University of Nottingham

School of Mathematical Sciences

Theoretical Neuroscience: From Long Wavelength Cortical Patterning to Spatial Navigation

Mayte Bonilla-Quintana, MSc

A thesis submitted to the University of Nottingham for the degree of

Doctor of Philosophy

June 2017

To my family and friends,
for their unconditional love,
support and encouragement.

Abstract

The science of brain function has a long and vibrant history. Recent technological developments have dramatically improved and facilitated data acquisition from a variety of methodologies to monitor brain activity, ranging from electroencephalography to opto-genetics. This highlights a need for concomitant theories of brain function. Such theories can act as a bridge between descriptions of the brain pertaining to data at different levels, from molecular to behavioural, using methods of mathematics, physics, and computer science.

The models presented in this thesis do not incorporate all the biophysical, anatomical and physiological data collected to date. Rather, the focus is on simplified models that contain sufficient detail to explain the essence of the phenomena considered. Moreover, they are constructed to allow the application of analytical mathematical tools to explore their behaviour.

In particular, this thesis proposes parsimonious neural models that aim to explain the mechanism by which humans and animals can navigate using spatial memory. The material presented ranges over a number of levels of description, and utilises a variety of mathematical techniques. A common theme throughout is the use of ideas from nonlinear dynamical systems to gain insight into neural mechanisms, ranging from activity patterns of cells underlying navigation, to the derivation of temporal difference reinforcement learning algorithms to solve reward based problems.

This work presents three main contributions. Firstly, it analytically determines which model parameters contribute to the observed difference in wavelength scale of the formed activity patterns in computational models for grid cells. Moreover, this thesis explores extensions to these models in order to find a neural mechanism that could account for the difference in wavelength scale. It is shown, after analysing the linear stability of spatially homogeneous steady

states to spatio-temporal perturbations, that the addition of axo-dendritic connections provides a mechanism for the difference in wavelength scale.

Secondly, based on recent research, this work proposes a different type of model, a network of spiking neurons, to uncover the mechanisms, related to rebound spiking, for variation in scale of grid cell firing fields. Travelling waves are observed on computer simulations of this model. The analytical construction of such waves is accomplished using techniques from the field of non-smooth dynamical systems. Moreover, the dispersion curve, that determines how wave speed varies as a function of the period, is constructed. Such dispersion curve exhibits a wide range of long wavelength solutions. In order to exhibit how the variation of parameters affects the maximum allowed period, a wave stability analysis is developed. This work entails and broadens the use of non-standard analysis techniques.

The final part of the thesis makes a direct link to experiments, combining reinforcement learning theory and computer simulations to shed light on the neurocomputational mechanisms underlying behaviour of rats in a variation of the Morris watermaze experiment. Particularly, the simulation employs a continuous time actor-critic framework, in which the actor and critic are represented as firing rate neural networks. The ability of the artificial rats to learn and reach the different goal locations is measured under different variations of the model.

Acknowledgements

First, I want to thank my supervisor, Steven Coombes, for his help and advice throughout these years. I am also very grateful to my co-supervisor, Reuben O'Dea, for all his support.

During my PhD I had the chance to collaborate with Robert Mason, Kyle Wedgwood, and Tobias Bast, to whom I want to thank for their time and guidance.

I am grateful to my sponsor in Mexico, Consejo Nacional de Ciencia y Teconología (CONACyT), and the University of Nottingham for their financial support.

I have been very fortunate to meet great people during my time in Nottingham that made this period of my life wonderful, to whom I am very thankful. Last but not least, I really appreciate all the love, help and encouragement from my family and friends - thank you for always been there for me.

Statement of originality

The content of this thesis is my own work. All the assistance received in preparing it and sources have been acknowledged. This work has not been submitted anywhere for any other degree or award.

Mayte Bonilla-Quintana

List of Abbreviations

CA	Cornu ammonis
Ca^{2+}	Calcium
CAN	Continuous Attractor Networks
Cl^{-}	Chloride
DMP	delayed-matching-to-place
EEG	Electroencephalography
HCN	Hyperpolarisation-activated cyclic-nucleotide-gated
HD	Head-Direction
HJB	Hamilton-Jacobi-Bellman
Hz	Hertz
I_h	hyperpolarisation-activated cation
K^{+}	Potassium
IF	Integrate-and-Fire
$\mathrm{Im}\,(z)$	Imaginary part of z
LIF	Leaky Integrate-and-Fire
MEC	Medial Entorhinal Cortex
MEC II	Layer II of the MEC
MPO	Membrane Potential Oscillations
mS	milliSiemens
mV	milliVolts
Na^{+}	Sodium
OIM	Oscillatory Interference Models
PIR	Postinhibitory Rebound
PWL	Piecewise Linear
$\mathrm{Re}\,(z)$	Real part of z
RBF	Radial Basis Function
RL	Reinforcement Learning
TD	Temporal Difference

Publications

The following paper has arisen from the work described in this thesis:

Mayte Bonilla-Quintana, Kyle CA Wedgwood, Reuben D. O'Dea, and Stephen Coombes. "An analysis of waves underlying grid cell firing in the medial enthorinal cortex." *The Journal of Mathematical Neuroscience* 7, no. 1 (2017): 9.

CONTENTS

1

INTRODUCTION

According to the Merriam-Webster dictionary, neuroscience is defined as the branch of the life sciences that deals with the anatomy, physiology, biochemistry or molecular biology of the nerves and nervous tissue, and especially with their relation to behaviour and learning [1]. In order to understand how the brain works, theories and technologies have been proposed with insights from other sciences, such as biology, physics, chemistry, psychology, and mathematics [2]. Neuroscience is therefore an intrinsically interdisciplinary field. Some of the quantitative approaches to neuroscience include: advanced data analysis, dynamical and statistical modelling, and theory (of neural function) [3]. Advanced data analysis consists of developing tools to obtain more understanding from neural data on multiple scales. These scales include the biophysical (cellular), the circuit, and the systems level [4]. The data obtained can be used to construct dynamical and statistical models that test the role of a specific mechanism. More sophisticated models can be developed in order to describe a theory to unify observations at different levels of description. Theories can also explain certain neural phenomena by identifying minimal structures or dynamics involved in that phenomena. Other theories see the brain as a computational unit and try to describe its function using algorithms. Theories about the brain are developed and studied in the field of *theoretical neuroscience*. On the other hand, *computational neuroscience* concerns data analysis, and the development of dynamical and statistical models. However, some computational neuroscientists study the brain as a computer [5], overlapping with the field of theoretical

neuroscience, and hence the terms computational and theoretical are sometimes used interchangeably.

The field of computational neuroscience started with the paper of Hodgkin and Huxley (see [6]) in which they describe the action potential in the giant axon of the squid, using nonlinear partial differential equations, fitted to electrophysiological data [7]. There has been a development of new mathematics inspired by the simplification of the equations presented in the Hodgkin and Huxley model, giving rise to the field of *mathematical neuroscience* [7]. For example, *neural fields* treat the interaction of billions of neurons as a continuum in order to provide a simpler description of a population with properties amenable to mathematical analysis [8]; although other theories emphasise need for electrophysiological properties in mathematical models [9].

The interpretation of the brain as a computer dates back to the seminal work of Marr [10]. His main interest was to understand how perception, feeling and thought are achieved by the brain. These phenomena can be understood as information-processing phenomenon. Marr identified three levels at which any machine carrying out an information processing task must be understood, namely, computational theory, representation[1] and algorithm, and hardware implementation. The computational theory contains separate arguments about what is computed and why, and it proposes an operation that is defined uniquely by the constraints it has to satisfy. Although Marr assumed that these three levels are separable, some authors do not make that assumption [11].

Recently, there has been an increased activity in the fields of theoretical and computational neuroscience, due to the technological improvement in the devices used to measure brain activity and in advances in computer hardware. Futhermore, the fields of theoretical and computational neuroscience have gained great support due the concern to avoid or replace, through mathematical and computer models, the use of animals in experiments [12].

1 A representation is a formal system for making explicit certain entities or types of information, together with a specification of how the system does this.

The aim of this thesis is to use the approach of mathematical neuroscience to gain knowledge about brain function, particularly navigation and spatial memory. A variety of perspectives are used to achieve this. For example, Chapter 3 examines how generic recurrently connected neural networks can generate spatio-temporal patterns. This is relevant to uncovering the mechanisms whereby grid cells can achieve different spatial scales of firing field patterns. Grid cells are spatially modulated neurons that fire at certain locations, known as firing fields, forming a hexagonal pattern. Such firing fields, and their varying scales, are known to be important for determining our sense of "space". In Chapter 3, the attention is restricted to rate models and various levels of biological reality are included, ranging from synaptic responses to axonal delays. The main method of analysis is that of linearisation and separation of variables.

Chapter 4 explores a different type of model, an integrate-and-fire model, in order to find an alternative mechanism for the different observed spatial scales of grid cells' firing patterns, by considering recent experimental discoveries of grid cells intrinsic properties. Computational simulations of the proposed model show the emergence of travelling waves whose orbits are subsequently mathematically described. Moreover, the Evans function for defining the stability of such travelling waves is derived using techniques of non-smooth dynamical systems. Importantly, the resulting mechanism for long wavelength pattern of this model is directly linked with experimental evidence.

A variation of the Morris watermaze task, the delayed-matching-to-place task, is simulated using an algorithm called reinforcement learning in Chapter 5. The aim of this simulation is to study how navigation towards a hidden location, without complete knowledge of the environment, can be achieved using neural rate models. Here, associations between the environment and neural activity are made and updated while exploring. The rules governing the update of such associations are derived.

This thesis starts by outlining the key theoretical concepts in Chapter 2 and concludes with some general remarks and future work in Chapter 6. The rele-

vant biological details and experimental evidence for each model are introduced throughout.

2

BACKGROUND

This chapter gives a brief description of the physiology and function of neurons, concentrating on the key components that are relevant to the models introduced and discussed in later chapters, and based on the work in [2, 4, 8].

A typical neuron consists of three main structures: the *dendritic tree*, the *cell body or soma*, and the *axon*. The *dendritic tree* processes the signals received from the *axons* of surrounding neurons in the form of an *electric potential*. The *electric potential* diffuses along the *dendritic tree* to the *soma* where it is processed. Alternatively, the contact of the axon to target neurons, also known as *synapses*, can be directly on the soma.

If the total potential at the soma exceeds a certain threshold value, the neuron produces a short electrical *spike* or *action potential* that is transmitted to the surrounding neurons via the axon, as shown in Figure 2.1. Both the dendritic tree and axon are branched structures so one neuron can receive electrical potential from, or send action potentials to, several thousand other neurons.

When an action potential arrives at a synapse, the presynaptic neuron activates voltage-dependent calcium (Ca^{2+}) channels producing an influx of Ca^{2+}. Such an influx leads to the secretion of a neurotransmitter that binds to receptors in the target (postsynaptic) neuron, causing ion-conducting channels to open. This process is described in the inset of Figure 2.1. Depending on the nature of the ion flow, the action potential arrival can either produce an increase in electrical potential (*excitatory synapse*) or a decrease (*inhibitory synapse*).

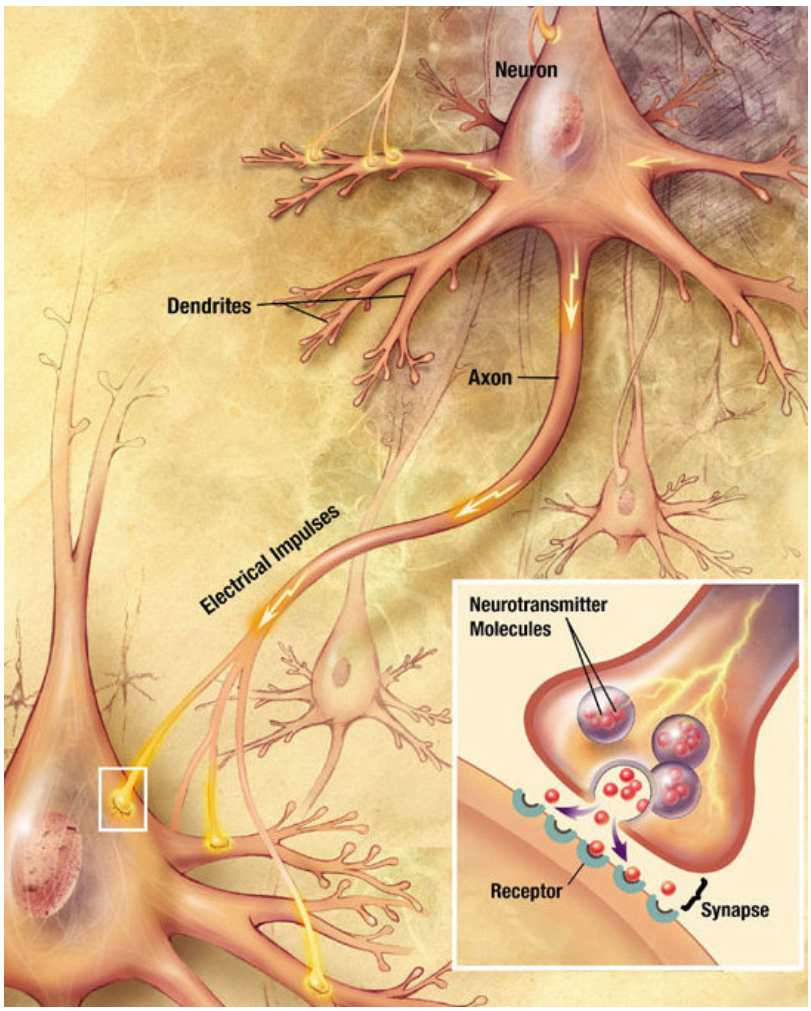


Figure 2.1. Schematic representation of the sypnatic process in neurons. Arrows denote the direction of the electrical impulses arriving to the dendrites from the axons after being processed in the soma. Figure obtained from https://commons.wikimedia.org/wiki/File%3AChemical_synapse_schema_cropped.jpg created by US National Institutes of Health, National Institute on Aging, via Wikimedia Commons

The cell is *polarised* when, under resting conditions, the potential inside the cell membrane is about -70 mV relative to the surrounding bath (0 mV). This potential difference is maintained by ion pumps located in the cell membrane that respond to changes in voltage and ion concentration. The cell *hyperpolarises* when positively charged ions flow out of the cell or negatively charged ions flow into the cell, making the membrane potential more negative. *Depolarisation* corresponds to (positive) ionic current flowing into the cell, making the membrane potential less negative. Thereby, the neuron generates an action potential when it is sufficiently depolarised to raise the membrane potential above a threshold level. After a cell spikes, a few milliseconds must pass before it can initiate another spike, this period is called the *absolute refractory period*. In fact, for a longer period, lasting tens of milliseconds, it remains difficult to generate an action potential. Such period is called the *relative refractory period*. Figure 2.2 illustrates the action potentials recorded at different locations of a neuron.

2.1 FIRING RATES

Figure 2.2 shows different types of action potentials (spike events). It is believed that action potentials convey information predominantly through their timing. In view of this, and their brief duration, they are often simply characterised by a spike sequence. Such a spike sequence can be represented as

$$\rho(t) = \sum_{i=1}^{n} \delta(t - t_i), \tag{2.1}$$

where t_i is the time of the ith spike, and δ is the Dirac-δ function that represents a infinitesimally narrow, idealised spike. Here $\rho(t)$ is called the *neural response function*.

Despite the trial-to-trial variability in spike sequences observed at a neuron, in most cases (since the number of synaptic inputs to most neurons is very large), the quantities of relevance to network dynamics are relatively insensitive to such fluctuations [4]. Therefore, it is often acceptable to use *firing rates* to

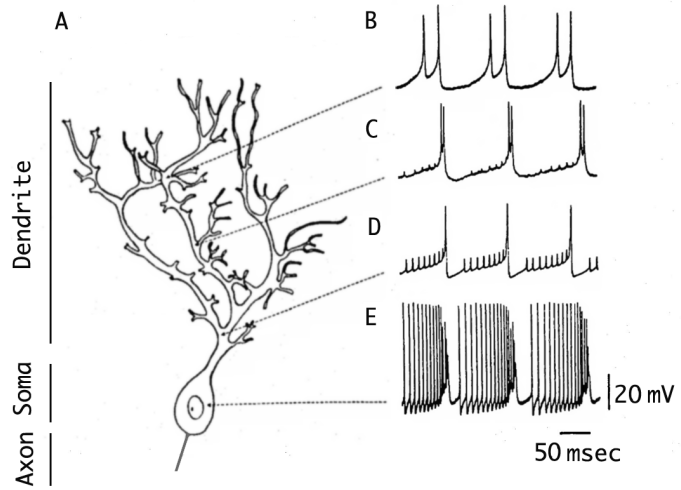


Figure 2.2. Action potentials (B-E) recorded at different locations of a neuron (A). The action potential illustrated in B shows an initial depolarisation followed by two spikes. After these spikes the cell hyperpolarises and then depolarises again. Note that the spike times and the subthreshold activity is different at different cell locations. Figure taken from [13].

characterise neural responses. A time-dependent firing rate can be constructed according to a time and ensemble average:

$$r(t) = \frac{1}{\Delta t}\int_{t}^{t+\Delta t} \mathrm{d}\tau \langle \rho(\tau)\rangle. \tag{2.2}$$

Here $\langle \rho(\tau)\rangle$ is the trial-average neural response function and $\Delta t > 0$ defines the window for temporal averaging. Note that $r(t)\Delta t$ is the probability that a spike occurs during the times t and $t + \Delta t$.

The response of a neuron to a given stimulus can be characterised by the number of action potentials fired during the presentation of such stimulus. Assuming that the stimulus is held constant over a trial, the *average firing rate* $\langle r\rangle$ is the average number of action potentials fired over trials and divided by the trial duration T:

$$\langle r\rangle = \frac{1}{T}\int_{0}^{T} \mathrm{d}t\; r(t). \tag{2.3}$$

This firing rate can be written as a function of a given stimulus s, i.e., $\langle r\rangle = f(s)$. The function $f(s)$ is known as the *neural response tuning curve*. The form of the tuning curves depends on the stimuli. For example, neurons in the primary

visual cortex of a monkey have Guassian-like response curves to a bar of light moved across the receptive field, as shown in Figure 2.3.

2.2 CELL DYNAMICS

The electrical signal of relevance in the nervous system is the difference in electrical potential between the interior and exterior of a neuron. Therefore the dynamics of a neuron can be described as an electrical circuit composed of resistors and capacitors, an example of such circuit is shown in Figure 2.4. Here, the cell membrane acts as a capacitor because it is impermeable to most charged molecules. The ion channels act as conductances (inverse of resistances) because they allow ions, predominantly sodium (Na^+), potassium (K^+), Ca^{2+} and chloride (Cl^-) to move into and out of the cell in response to voltage changes (voltage dependent channels) and both internal (Ca^{2+}-dependent channels) and external signals (synaptic receptor channels); thus lowering the effective membrane resistance to ion flow. The channels are further categorised into *passive*, *active* and *ligand-gated*. *Passive* signifies that the channel conductance does not depend on the cell's membrane potential, whereas *active* signifies that it does. A *ligand-gated* channel requires the binding of a helper molecule (the ligand) to open the channel [2]. The resultant membrane conductance depends on the density and type of channels.

The amount of current required to change the membrane potential at a given rate can be determined by using the standard equation for a capacitor that relates voltage and charge as:

$$Q = C_m V, \tag{2.4}$$

where C_m is the capacitance, V the voltage across the membrane and Q the amount of excess negative charge on the inside surface of the neuron's cell membrane. Differentiation with respect to t gives the evolution equation

$$C_m \frac{\mathrm{d}V}{\mathrm{d}t} = \frac{\mathrm{d}Q}{\mathrm{d}t}. \tag{2.5}$$

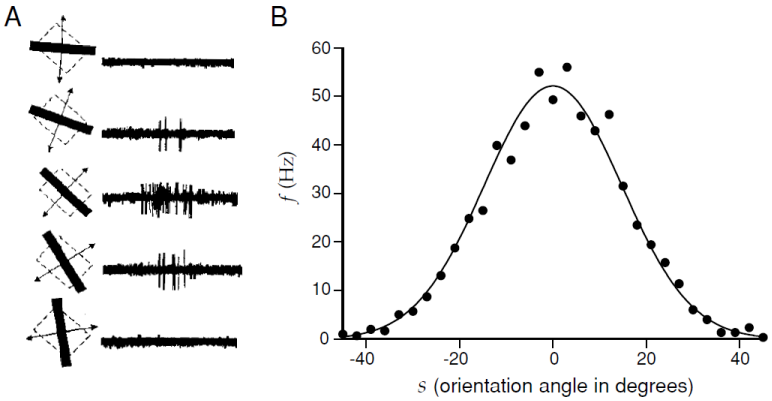


Figure 2.3. (A) Right: Activity recorded from a neuron in the primary visual cortex of a monkey. Left: Bar of light (black bar) moved across the receptive field of the cell (square dotted box) at different angles. The motion of the light is indicated by the arrows. (B) Average firing rate of a cat neuron in the visual cortex plotted as a function of the orientation angle of the light bar stimulus. Figure taken from [4].

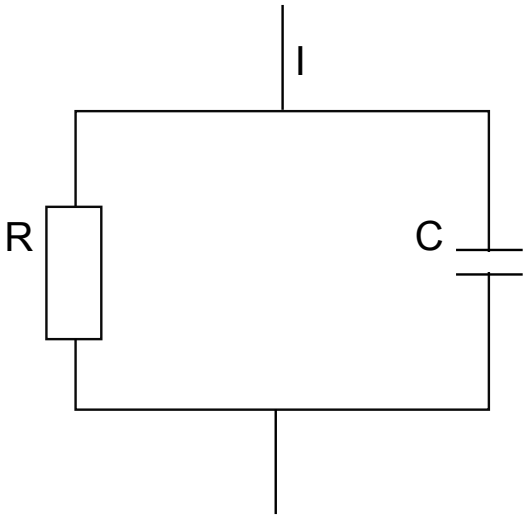


Figure 2.4. Diagram of a typical resistor-capacitor circuit depicting a neuron. Here the capacitor (C) and the resistor (R) lie in parallel.

The membrane potential can be shifted away from its resting value by the injection of a small constant current. Denoting the shift in potential by ΔV and the injected current by I_e, Ohm's law gives:

$$\Delta V = I_e R_m, \tag{2.6}$$

where R_m is the membrane resistance.

When the neuron is inactive, the flow of ions into the cell matches that out the cell, thereby the potential is at an equilibrium point. The potential changes when the balance of ion flow is modified by the opening or closing of ion channels. Following [14], and assuming that the potential is at equilibrium, the probability P_{in} of finding a specific ion inside the cell can be compared to the probability P_{out} of finding it outside (using concepts of statistical mechanics). Such a comparison is given by the Boltzmann distribution

$$\frac{P_{in}}{P_{out}} = \frac{\exp\left(-\frac{E_{in}}{k_B \Gamma}\right)}{\exp\left(-\frac{E_{out}}{k_B \Gamma}\right)}, \tag{2.7}$$

that depends on the energy the ion has inside (E_{in}) and outside (E_{out}) the cell. Here k_B is a constant called the Boltzmann factor and Γ is the temperature (in Kelvin). The difference between the energy of the ions inside or outside the cell is caused by the difference in the electrical energy. This energy is given by zeV, where z is the valence of the ion, e the electronic unit charge, and V the electric potential. Therefore

$$\frac{P_{in}}{P_{out}} = \exp\left(-\frac{ze\left(V_{in} - V_{out}\right)}{k_B \Gamma}\right). \tag{2.8}$$

The voltage difference between the inside and outside of the cell can be obtained by manipulating (2.8), yielding

$$E = V_{in} - V_{out} = \frac{k_B \Gamma}{ze} \ln \frac{P_{out}}{P_{in}}, \tag{2.9}$$

in which E is commonly called the *reversal potential* and conventionally $V_{out} = 0$. The above equation is known as the Nernst equation and relates the probability of finding an ion at some location to its concentration at this location. The conductances tend to move the membrane potential of the neuron towards

its reversal potential. When $V > E$, positive current will flow outward, whereas when $V < E$ positive current will flow inward. If conductances have reversal potentials near to the resting potential then they pass little net current and their primary impact is to change the membrane resistance.

Following [4], the total membrane current per unit area of cell membrane, i_m, is given by

$$i_m = \sum_i g_i(V - E_i), \tag{2.10}$$

where E_i is the reversal potential of channel type i, $V - E_i$ is the driving force, and g_i is the conductance per unit area.

The rate at which charge builds up ($\mathrm{d}Q/\mathrm{d}t$ in (2.5)) is equal to the amount of current entering the neuron (given by the sum of (2.6) and (2.10)). Therefore

$$c_m \frac{\mathrm{d}V}{\mathrm{d}t} = -i_m + \frac{I_e}{A} \tag{2.11}$$

where A is the total surface area of the neuron, $c_m = C_m/A$ is the specific membrane capacitance and i_m is, by convention, positive when positive ions leave the neuron and negative when positive ions enter the neuron.

If all the active conductances are ignored and the total membrane conductance is modelled as a single passive leakage term, $i_m = \overline{g}_L(V - E_L)$ with $\overline{g}_L$ a constant, then the *passive compartment model* for neuronal dynamics is given by

$$\tau_m \frac{\mathrm{d}V}{\mathrm{d}t} = E_L - V + R_m I_e, \tag{2.12}$$

where $R_m = r_m/A$ is the total membrane resistance and $\tau_m = c_m r_m$ the membrane time constant. Here $r_m = 1/\overline{g}_L$ is the specific membrane resistance.

2.2.1 *Active Membrane*

Membrane conductances change over time and most of the interesting electrical properties of neurons arise from non-linearities associated with active membrane conductances [4]. Experimental evidence shows that ion channels fluctuate rapidly between open and closed states. Because there is a large number of

channels and they fluctuate independently of each other, the fraction of channels open at a given time can be approximated with the probability that any channel is in an open state. Therefore, the value of the conductance per unit area of membrane due to a set of channels of type i, g_i, at any given time is determined by

$$g_i = \overline{g}_i P_i, \tag{2.13}$$

where $\overline{g}_i$ is called the maximal conductance (given by the product of the conductance of an open channel by the density of the channels) and P_i is the probability of finding any given channel in an open state.

2.2.1.1 *Voltage-Dependent Conductance*

The ion channels that produce the so-called *persistent conductance* can be thought of as pores with a gate that opens or closes depending on the voltage across the membrane. When the gate opens, it *activates* the conductance, whereas when it closes, the conductance is *deactivated*. An example of a persistent conductance is the delayed-rectifier K^+ conductance that is responsible for repolarising a neuron after an action potential. For this type of conductance the probability that the gate is open, P_K, increases when the neuron is depolarised and decreases when it is hyperpolarised. Moreover, the gate opens when k identical subgates open, thus

$$P_K = n^k, \tag{2.14}$$

where n is called the *gating variable* and is the probability that any one of the k independent subgates is in an open state. The gating variable changes according to a law of mass action

$$\frac{\mathrm{d}n}{\mathrm{d}t} = \alpha_n(V)(1-n) - \beta_n(V)n, \tag{2.15}$$

where $\alpha_n(V)$ is the opening rate and $\beta_n(V)$ is the closing rate. Usually (2.15) is written as

$$\tau_n(V)\frac{\mathrm{d}n}{\mathrm{d}t} = n_\infty(V) - n, \tag{2.16}$$

where $\tau_n(V)$ is the time constant

$$\tau_n(V) = \frac{1}{\alpha_n(V) + \beta_n(V)}, \tag{2.17}$$

and

$$n_\infty(V) = \frac{\alpha_n(V)}{\alpha_n(V) + \beta_n(V)} \tag{2.18}$$

describes the steady-state levels of activation of the K^+ conductance.

There is another type of channel that opens transiently and is controlled by two gates with opposite voltage dependences. The conductance generated by such channels is therefore called *transient*. One example is the fast open Na^+ conductance with

$$P_{Na} = m^k h, \tag{2.19}$$

where m is the activation variable and h is the inactivation variable. Whilst m behaves like n in (2.14), h decreases when the neuron is depolarised and increases when it is hyperpolarised. Thus, m and h are described by equations identical to (2.16) with α_m, α_h and β_m, β_h fitted accordingly. Because h is an inactivation variable, it approaches 1 at hyperpolarised voltages and 0 at depolarised voltages, and therefore $h_\infty(V)$ is flipped relative to $m_\infty(V)$ and $n_\infty(V)$. All gating variables in modern conductance based models of active membrane are described by equations similar to (2.16). The rate functions $\alpha_j = \alpha_j(V)$ and $\beta_j = \beta_j(V)$ can be fitted to data, as in the famous Hodgkin and Huxley model [6].

A *hyperpolarisation-activated* conductance is controlled only by an inactivation gate that opens when the neuron is hyperpolarised. This type of conductance is studied in Chapter 4.

2.2.1.2 *Synaptic Conductance*

The probability of an open channel P_{syn} for a synaptic conductance depends on the probability P_s that a postsynaptic channel opens (or fraction of channels opened), given that neuro transmitter was released by the presynaptic terminal,

and the probability P_{rel} that the presynaptic terminal released the transmitter after the arrival of an action potential. Therefore, P_{syn} is given by

$$P_{syn} = P_s P_{rel}. \tag{2.20}$$

The probability that the channel is open changes over time according to

$$\frac{\mathrm{d}P_s}{\mathrm{d}t} = \alpha_s(1 - P_s) - \beta_s P_s. \tag{2.21}$$

Here, the closing rate of the channel, β_s, is usually assumed to be constant whereas the opening rate α_s depends on the concentration of transmitter available for binding to the receptor.

One common way to describe a postsynaptic conductance for an isolated presynaptic release that occurs at time $t = 0$, reaches its peak at $t = 1/\alpha$, and after which decays with time constant $1/\alpha$, is to use the so-called alpha function given by

$$\eta(t) = \alpha^2 t \exp(-\alpha t) H(t), \tag{2.22}$$

where $H(t)$ is the Heaviside step function defined by

$$H(t) = \begin{cases} 1, \text{ if } t \geq 0 \\ 0, \text{ otherwise} \end{cases}. \tag{2.23}$$

Thus, the postsynaptic conductance change is given by

$$g(t) = \overline{g}\eta(t - T), \tag{2.24}$$

where T is the arrival time of a pre-synaptic action potential. The change arising from a train of action potentials, with firing times t_m, can hence be written as

$$g(t) = \overline{g} \sum_m \eta(t - t_m). \tag{2.25}$$

Following [8], the above equation can be written as the *Green's function* of a linear differential operator[1], Q, so that $Q\eta = \delta$, where

$$Q = \left(1 + \frac{1}{\alpha}\frac{\mathrm{d}}{\mathrm{d}t}\right)^2. \tag{2.26}$$

Therefore, an equivalent form for (2.25) is

$$Qg = \overline{g}\sum_m \delta(t - t_m) = \overline{g}\rho(t), \tag{2.27}$$

where $\rho(t)$ is the neural response function in equation (2.1).

2.2.2 *Quasi-Active Membrane*

A passive membrane model (2.12) cannot accomplish certain kinds of specific neural dynamics. For example, *resonances*, which describe the increase in the likelihood of firing when the neurons have inputs at specific preferential frequencies. In order to exhibit these specific neural dynamics, the model has to include more information relating to the ion channel activity. Interestingly, Coombes *et al.* [15] showed that a linearisation of the channel kinetics about rest may describe adequately such cell dynamics, simplifying the Hodgkin and Huxley type model described in §2.2.1.1. Following their work, a generic ionic membrane current of the form

$$I = I(V, w_1, \ldots, w_N), \tag{2.28}$$

is considered, where V is a voltage and w_k are gating variables that satisfy (2.16). Thus, the steady state of the current is given by $\overline{I} = I(\overline{V}, w_{1,\infty}(\overline{V}), \ldots, w_{N,\infty}(\overline{V}))$. Taking small perturbations about the steady state of the form

$$(V, w_1, \ldots, w_N) = (\overline{V} + \delta V, w_{1,\infty}(\overline{V}) + \delta w_1, \ldots, w_{N,\infty}(\overline{V}) + \delta w_N), \tag{2.29}$$

1 In general, the Green's function $G = G(t, s)$ at a point s corresponding to a linear operator $\mathcal{L}$ is any solution of $\mathcal{L}G(t, s) = \delta(t - s)$, where δ denotes the Dirac-delta function. This identity is used to solve differential equations of the form $\mathcal{L}u(t) = f(t)$, by multiplying $\mathcal{L}G(t, s) = \delta(t - s)$ by the function $f(s)$ and integrating with respect to s. Thus, the solution of the differential equation is given by $u(t) = \int G(t, s)f(s)\mathrm{d}s$.

and substituting in (2.28) yields

$$\delta I = \frac{\delta V}{R} + \sum_{k=1}^{N} \left.\frac{\partial I}{\partial w_k}\right|_{ss} \delta w_k, \tag{2.30}$$

where the resistance R is defined by $R^{-1} = \partial I/\partial V|_{ss}$ and $_{ss}$ refers to the steady state. Using (2.15) the evolution of the perturbations in the gating variable can be written as

$$\left(\frac{\partial}{\partial t} + \alpha_k + \beta_k\right) \delta w_k = \left[\frac{\mathrm{d}\alpha_k}{\mathrm{d}V} - w_{k,\infty} \frac{\partial\left(\alpha_k + \beta_k\right)}{\partial V}\right] \delta V. \tag{2.31}$$

Substituting (2.31) in (2.30) gives

$$\delta I = \frac{\delta V}{R} + \sum_{k=1}^{N} \delta I_k, \tag{2.32}$$

where

$$\left(r_k + L_k \frac{\mathrm{d}}{\mathrm{d}t}\right) \delta I_k = \delta V, \tag{2.33}$$

with

$$r_k^{-1} = \tau_k \frac{\partial I}{\partial w_k} \left[\frac{\partial \alpha_k}{\partial V} - w_{k,\infty} \frac{\partial\left(\alpha_k + \beta_k\right)}{\partial V}\right]\Bigg|_{ss}, \tag{2.34}$$

$$L_k = \tau_k r_k. \tag{2.35}$$

Now, considering a general current balance equation of the form

$$C\frac{\mathrm{d}V}{\mathrm{d}t} = -g_L(V - V_L) - I(V, w_1, \ldots, w_N), \tag{2.36}$$

the following linearised equations are obtained

$$C\frac{\mathrm{d}V}{\mathrm{d}t} = -\frac{V}{\overline{R}} - \sum_{k=1}^{N} I_k, \qquad \frac{1}{\overline{R}} = g_L + \frac{1}{R}, \tag{2.37a}$$

$$L_k \frac{\mathrm{d}I_k}{\mathrm{d}t} = -r_k I_k + V. \tag{2.37b}$$

This describes an "LCR" circuit (see Fig. 2.5), in which the current I responds as though the resistance R is in parallel with N impedance lines. Each of these is a resistance r_k that is itself in series with an inductance L_k. In §3.5 a quasi-active model of cell dynamics is proposed in order to account for the resonant behaviour seen experimentally in grid cells.

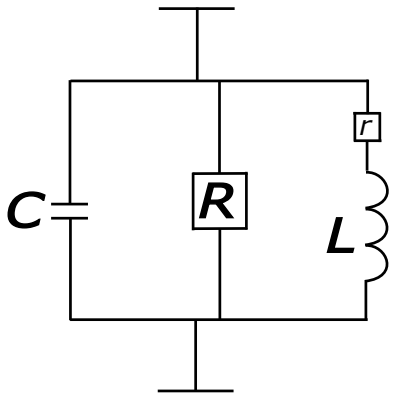


Figure 2.5. An "LCR" circuit depicting a quasi-active membrane model (2.37), in which the membrane acts as a capacitor that is in parallel with a resistance R and an inductance L.

2.3 DENDRITES

Membrane potentials measured in the soma, dendrite and axon of a cell can take different values as shown in Figure 2.2. Potential differences between different parts of a neuron cause ions to flow within the cell, which tends to equalise these differences. The intracellular medium provides a resistance to such a flow. This resistance is highest for long, narrow stretches of dendritic or axonal cable. Following [8], the conservation of electrical current in such a cylindrical element of a neuron may be described by the so-called *cable equation*.

Let $V(\xi, t)$ denote the membrane potential at position ξ along an infinite uniform cable at time t measured relative to the resting potential of the membrane. Let τ be the cell membrane time constant, λ the space constant, and r the membrane resistance, then the uniform (infinite) cable equation is

$$\tau \frac{\partial V(\xi,t)}{\partial t} = -V(\xi,t) + \lambda^2 \frac{\partial^2 V(\xi,t)}{\partial \xi^2} + rI(\xi,t), \quad \xi \in (-\infty, \infty), \tag{2.38}$$

where the source term $I(\xi, t)$ corresponding to external input injected into the cable is included. Let $I(\xi, t)$ be the response to a unit impulse at ξ' at $t = 0$, and take $V(\xi, 0) = 0$, then the dendritic potential at the point ξ behaves as

$$V(\xi,t) = rG_\infty(\xi - \xi', t)/\tau, \tag{2.39}$$

with

$$G_\infty(\xi,t) = \frac{1}{\sqrt{4\pi D t}} \mathrm{e}^{-t/\tau} \mathrm{e}^{-\xi^2/(4Dt)} H(t). \tag{2.40}$$

Here $D = \lambda^2/\tau$ is the diffusion constant and (2.40) is the Green's function for the relevant operator. Note that the Green's function determines the linear response to the unit impulse at ξ' at $t = 0$.

Due to the linearity of (2.38), the solution for any given input can be constructed as a superposition of (2.40). Therefore the general solution of (2.38) is given by

$$\begin{aligned} V(\xi,t) &= \frac{r}{\tau}\Bigg(\int_{-\infty}^{t} \mathrm{d}t' \int_{-\infty}^{\infty} \mathrm{d}\xi' \, G_\infty(\xi-\xi', t-t') I(\xi', t') \\ &+ \int_{-\infty}^{\infty} \mathrm{d}\xi' \, G_\infty(\xi-\xi', t) V(\xi', 0) \Bigg). \end{aligned} \tag{2.41}$$

The full details are shown in Appendix A. Note that the spatio-temporal distribution of delays, expressed by the Green's function in (2.41), is generated by the diffusion along the dendritic tree. The dynamics induced by incorporating dendritic processing to a neural network are studied in §3.6.

2.4 OSCILLATORY INTERFERENCE MODELS

Some of the neural models proposed in this thesis are compared with models in the literature, for example, the oscillatory interference models. Thus a brief introduction to this type of model is given here.

The voltage fluctuations of the brain over time can be recorded using electroencephalography (EEG), in which electrodes are placed along the scalp or directly on the brain (via an invasive surgery). Sometimes the recorded activity can be irregular or exhibit oscillations; for example, the hippocampal EEG of a rat can exhibit *theta* (sinusoidal 7-12 Hz waves), or highly irregular activity over a larger frequency domain [16]. The type of activity exhibited in the EEG depends on the behaviour of the rat. When it is walking, exploring, running, swimming, rearing, and jumping the EEG has theta activity; whereas, when the rat is doing other activities that do not require any displacement movement,

the EEG activity is irregular. Interestingly, the activity from a single neuron exhibits theta only when the rat is near to a particular location z. O'Keefe and Recce proposed a model for this phenomenon [16]. In their model a neuron is an oscillator that receives input from at least two different rhythmic generators at slightly different frequencies. Thus, throughout most of the environment these generators cancel each other, but at the location z the frequency of one of the oscillators increases slightly relative to the other. Assuming that the rat is moving at a constant speed and direction, the activity y of the neuron is given by

$$y = 2a\cos\left[0.5(k_1 + k_2)x - 0.5(\omega_1 + \omega_2)t\right] \cdot \sin\left[0.5(k_1 - k_2)x - 0.5(\omega_1 - \omega_2)t\right], \tag{2.42}$$

where a is a constant, x represents the position of the rat at time t, k is the spatial propagation number of each wave and ω is the angular frequency of the temporal component. Modified versions of this model have been used to represent the activity of some type of neurons used in navigation (see [17] for a review).

2.5 NETWORK MODELS

2.5.1 *Neural Field Models*

In order to construct a network model, the total synaptic input received by a neuron, as a result of anatomical connectivity, is considered. The total synaptic input is defined by Dayan and Abbot [4] as the total current delivered to the soma as a result of all conductance changes resulting from presynaptic action potentials.

Let us define the total synaptic current as $I_{syn} = g_{syn}(V - V_{syn})$, where V_{syn} is a reversal potential and g_{syn} a conductance as defined in (2.27). Then,

assuming (as in [8]) that a neuron spends most of its time close to rest, such that $V_{syn} - V \approx V_{syn}$, and absorbing the factor V_{syn} in g_{syn}, yields

$$I_{syn} \simeq g_{syn} = \overline{g} \sum_m \int_{-\infty}^{t} \mathrm{d}s \eta(t-s)\delta(s-t_m). \tag{2.43}$$

Using an alternative representation with the Green's function $\eta(t)$ (2.43) can be rewritten as

$$Qg_{syn} = \overline{g} \sum_m \delta(t-t_m). \tag{2.44}$$

Due to the complexity of simulating, analysing and interpreting networks of spiking neurons, the outputs of neurons are usually modelled as firing rates. These firing rate models avoid the short time scale dynamics of action potentials and model a slower collective behaviour. Here, η is assumed to be sufficiently slow (i.e. $1/\alpha \to 0$) so that the short-time average, defined by

$$\langle x \rangle_t = \frac{1}{\Delta t} \int_{t-\Delta t}^{t} x(s)\mathrm{d}s, \tag{2.45}$$

of the left hand side of (2.44) is approximately constant. Applying (2.45) to the right hand side of (2.44) gives the so-called *instantaneous firing rate function*, f. This function is assumed to be a function of the drive alone for a single neuron experiencing a constant drive. For synaptically interacting neurons, the drive is directly proportional to the conductance state of the presynaptic neuron, thus $f = f(g_{syn})$. Therefore, a model for a single neural population with self-feedback is given by

$$Qg_{syn} = wf(g_{syn}), \tag{2.46}$$

where w represents the strength of coupling, and for excitatory synapses $w > 0$ whereas for inhibitory, $w < 0$.

Now consider the case that neurons are arranged in a tissue where the anatomical connectivity between two neurons, with spatial coordinate $x, y \in D$ respectively, is $w(x,y)$. Then the total synaptic input to the neuron at the spatial coordinate x at time t is given by

$$I_{syn}(x,t) \simeq g_{syn}(x,t) = \int_{-\infty}^{t} \mathrm{d}s \eta(t-s) \int_D \mathrm{d}y \; w(x,y) f(g_{syn}(y,s)). \tag{2.47}$$

The synaptic input plays a major role in the network models presented in this thesis. Chapters 3 and 4 study the emergent activity from a network with inhibitory connectivity. In the former, rate functions are used to describe the dynamics of synaptic currents whereas in the later spike trains are used. In Chapter 5 plasticity is considered, in which the synaptic weights can also evolve dynamically.

2.5.2 *Continuous Attractor Network Models*

In order to represent a group of neurons with the same characteristics, namely a subpopulation or pool of neurons, some computational models use *nodes* [18]. A node receives some input r_i^{in} from other nodes, does some local processing and distributes the generated output r^{out} to other nodes. Here $r_i^{in}, r^{out} \in \mathbb{R}$, although binary variables may also be used. The most basic type of node is called a *sigma node*. This type of node has several inputs r_i^{in} that it processes to generate only one output r^{out} as follows

$$r^{out} = \psi\left(h\right) = \psi\left(\sum_i w_i r_i^{in}\right), \tag{2.48}$$

where ψ is known as the activation or gain function and w_i is the weight of the input i. The node sums up all the weighted inputs and the resulting value $h = \sum_i w_i r_i^{in}$ is known as the activation of the node. Note that (2.48) describes the single node response to an external input, not considering its evolution over time. In order to describe the time dynamics of a sigma node, the following assumptions are needed: 1) the processing done by the sigma node in (2.48) takes a fixed time length τ; 2) the continuous node integrates only a fraction $\Delta t/\tau$ of the input $\sum_i w_i r_i^{in}$ in each time step $\Delta t < \tau$; and 3) the node has to remember its previous values to which the new input is added. Hence,

$$h(t + \Delta t) = h(t) + \frac{\Delta t}{\tau} \sum_i w_i r_i^{in}. \tag{2.49}$$

Note that h will continue to grow without bounds as long as the input is applied[2]. In [18], this problem is solved by gradually leaking a small amount of the input that the node has integrated before. This small amount is called the forgetting factor. Now, if the forgetting factor is set on the same time scale as the integration of new input in (2.49), then

$$h(t+\Delta t) = \left(1 - \frac{\Delta t}{\tau}\right) h(t) + \frac{\Delta t}{\tau} \sum_i w_i r_i, \tag{2.50}$$

Note that if $\Delta t = \tau$ then (2.48) is obtained. After taking the limit $\Delta t \to 0$ (2.50) gives the temporal dynamics of the activation of a sigma node

$$\tau \frac{\mathrm{d}h(t)}{\mathrm{d}t} = -h(t) + \sum_i w_i r_i; \tag{2.51}$$

which has the following analytical solution

$$h(t) = h(0)\mathrm{e}^{\frac{-t}{\tau}} + \sum_i w_i r_i \left(1 - \mathrm{e}^{\frac{-t}{\tau}}\right). \tag{2.52}$$

Note that if $\sum_i w_i r_i = 0$, the activity of the node just decays exponentially in time. On the other hand if the input is positive, it slows down the exponential decay of h; and if the input is negative, it decreases the activation of the node. But in all cases, $h \to \sum_i w_i r_i$ as $t \to \infty$. Thus, regardless of the input value, setting the forgetting factor on $\Delta t/\tau$ ensures that the activation of the node will settle to a define value, guaranteeing its stability.

Now, consider a network of N sigma nodes, where h_i represents the activation of the ith node. Such a network is recurrently connected and the strength of input from node j to node i is represented by w_{ij}. The dynamics of this network is given by

$$\frac{1}{\tau}\frac{\mathrm{d}h_i(t)}{\mathrm{d}t} = -h_i(t) + \frac{1}{N}\sum_j w_{ij} r_j(t) + I_{ext_i}(t), \tag{2.53}$$

where $I_{ext_i}(t)$ is an external input to the system, τ is a time constant and $r_j(t) = \psi(h_j(t))$. The dynamics of such a system evolve in time and eventually settle down to a single stable equilibrium state, a collection of equilibrium states, or a periodic oscillation, depending on the initial conditions. In computational

2 The solution of (2.49) in the limit when $\Delta t \to 0$ is $h(t) = \frac{\sum_i w_i r_i^{in}}{\tau} t + h(0)$.

neuroscience these models are known as *attractor networks* and there are many types depending on the attractor states. These attractor states are used to represent the "memory" of the system, in the sense that the system will be able to retrieve such a memory when noise in the system or in the connections is present (i.e., the system evolves to the attractor state). For example, the *continuous attractor* allows a continuum of stable equilibrium states, or at least closely related states that approximately form a continuum, and are well-suited to represent two-dimensional variables, such as location in a room [19]. Thus, continuous attractor network (CAN) models have been used to model parts of the brain related to spatial memory (see [19] for a review). One such model is studied in Chapter 3.

In attractor network models, the attractor states refer to states of activity in the network. Generally, there is no anatomical correspondence between attractor states in the network and the physical locations in brain [19]. For example, in 1D CAN models of head-direction (HD) cells[3], nodes are conceptualised as being arranged in a circle (or ring), where the location of each node in the circle with respect of the circle's centre represents the preferred firing direction of one or a pool of HD cells as shown in Figure 2.6. Nearby nodes code for similar preferred directions although experimental evidence shows that nearby HD cells have different preferred firing directions [20]. Particularly, in the model proposed by McNaughton *et al.* [21], neighboring nodes are connected by strong excitatory synapses and nodes that are far apart (in the circle) are connected with inhibitory synapses. The strength of these connections is proportional to the angular distance between the nodes on the circle. This symmetric connectivity between nodes results in a "bump" of activity on the ring (attractor state), because the excitatory circuitry reinforces the activity of other nearby nodes, whereas the inhibitory connections force nodes further away to be silent. It is

3 Head-direction cells fire whenever a subject's head points in a particular direction in an allocentric coordinate frame [19]. Different head-direction cells fire at different preferred directions. An allocentric coordinate is based on the external environment (independent of the subject's location). Thus, in this frame, the subject's head points to the north, east, etc.

important to note that the heading of the animal changes when it is wandering an environment. Therefore, the location of the activity bump has to change accordingly to the node that encodes for that particular heading direction. In order to achieve this, an asymmetric input modulated by the rat's velocity and heading direction is added to the HD nodes. Such input is higher to the nodes whose preferred heading direction is similar to the animal's heading direction. The resulting model is a CAN model. It is important to note that node location represents the preferred head direction and not physical location of the cell. Despite this fact CAN models of HD cells can explain a number of properties of the HD system such as the firing of HD cells in the same preferred direction across multiple environments and the preservation of cell preferred firing in the dark [19].

2.6 SUMMARY

This chapter gave a short description of different types of models that are commonly used to study neural phenomena, from single neuron activity fluctuations due to the influx of ions, through to tissue activity arising mainly from the connections between neurons. The type of model that will be implemented in the following chapters will depend on the phenomena described in each one; for example, in Chapter 4 the microscopic detail of ion channel fluctuations is included, whereas Chapter 5 studies the overall activity of different groups of neurons.

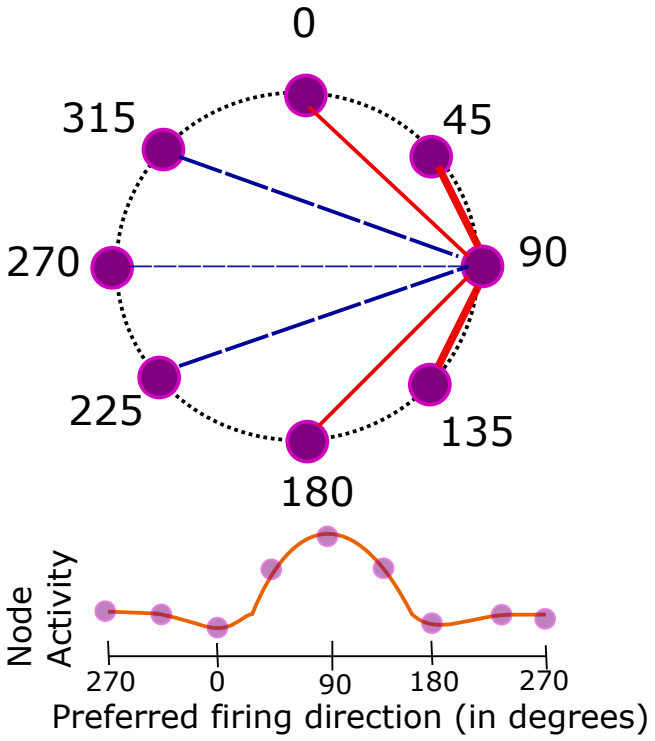


Figure 2.6. Model of Head-Direction cells. Top: Purple circles represent the nodes arranged on a black dotted circle. Note that each node codes for a different preferred firing direction that is displayed next to the node. Red lines denote excitatory connections and blue dashed lines denote inhibitory connections. The thickness of these lines represents the strength of the connections. For simplicity only the connections of the node coding for a preferred direction of 90° are shown. Bottom: Orange line represents the bump of activity formed when the subject head points to east (90°) direction.

3

MECHANISMS FOR LONG WAVELENGTH PATTERNING IN FIRING RATE NEURAL NETWORKS

3.1 INTRODUCTION

The aim of this chapter is to investigate the classes of neural network models that can support long wavelength spatial patterns. Such patterns can be of physiological significance in a wide range of settings. Here, the motivation is to understand the mechanisms that can underlie the so-called grid cell firing.

Grid cells are neurons in the Medial Entorhinal Cortex (MEC), that fire at multiple locations when an animal is wandering in an open two-dimensional environment, defining a periodic triangular array that covers the entire surface of the environment making a grid-like pattern, hence the name [22] (Fig. 3.1). These cells fire at the same position regardless of changes in the animal's speed and direction, and firing persists in the absence of visual input. Grid cells are therefore believed to correspond to the animal's own sense of location. For their discovery in 2005 [22], May-Britt and Edvard Moser were awarded the Nobel Prize in Physiology and Medicine in 2014 (together with John O'Keefe).

The MEC is organised in layers that are embedded one above the other and have different thickness. Each of these layers are characterised by a dominant type of cell and by their principal inputs and/or outputs within the layer and to other brain regions. Although grid cells can be found in all MEC layers, experiments in rats done by Boccara *et al.* [23] showed that the number of grid cells decreases from superficial layers (II and III) to deeper layers (V and VI) of

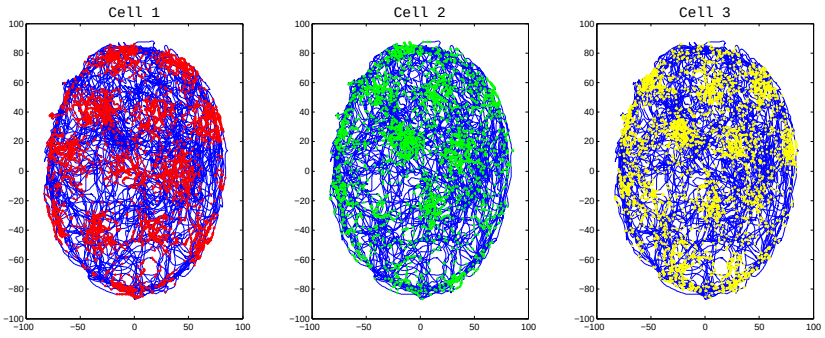


Figure 3.1. Trajectory of a rat (blue line) running on a cylinder with a diameter of 180 cm. Red, green and yellow dots correspond to the locations of the spikes of three different cells recorded simultaneously in layer II at the dorsocaudal pole of the Medial Entorhinal Cortex. Experimental data was obtained by Hafting and it is available on the Moser Group website http://www.ntnu.no/cbm/moser.

the MEC. Boccara *et al.* also found grid cells in pre- and parasubiculum, but their proportion was comparable to deep layers of MEC. The MEC provides input to the *hippocampal formation*, that consists of dentate gyrus, subiculum, CA1, CA2 and CA3; as shown in Figure 3.2. Particularly, layer II of MEC projects to the dentate gyrus, CA3 and CA2, and layer III projects to CA1 and the subiculum [24]. The principal neurons of the Entorhinal Cortex (EC)[1] are generally pyramidal cells or modified versions, the so-called stellate cells [25]. Stellate cells form the large majority of principal neurons in MEC II. Couey *et al.* [26] showed that MEC II stellate cells are mainly interconnected via inhibitory interneurons (with short axons) and that excitatory connections between stellate cells are essentially absent.

The hippocampal formation along with the presubiculum, parasubiculum, and entorhinal cortex are believed to provide a metric representation of the environment [27]. There are other spatially modulated cell types in the hippocampal formation besides grid cells, namely *place cells*, *head direction cells*, *boundary cells* and cells that encode object location. Place cells are found in the CA1 and fire selectively to spatial locations, whereas head direction cells

1 Here principal neurons refers to the neurons that receive most of the incoming axons and are the major output source of output.

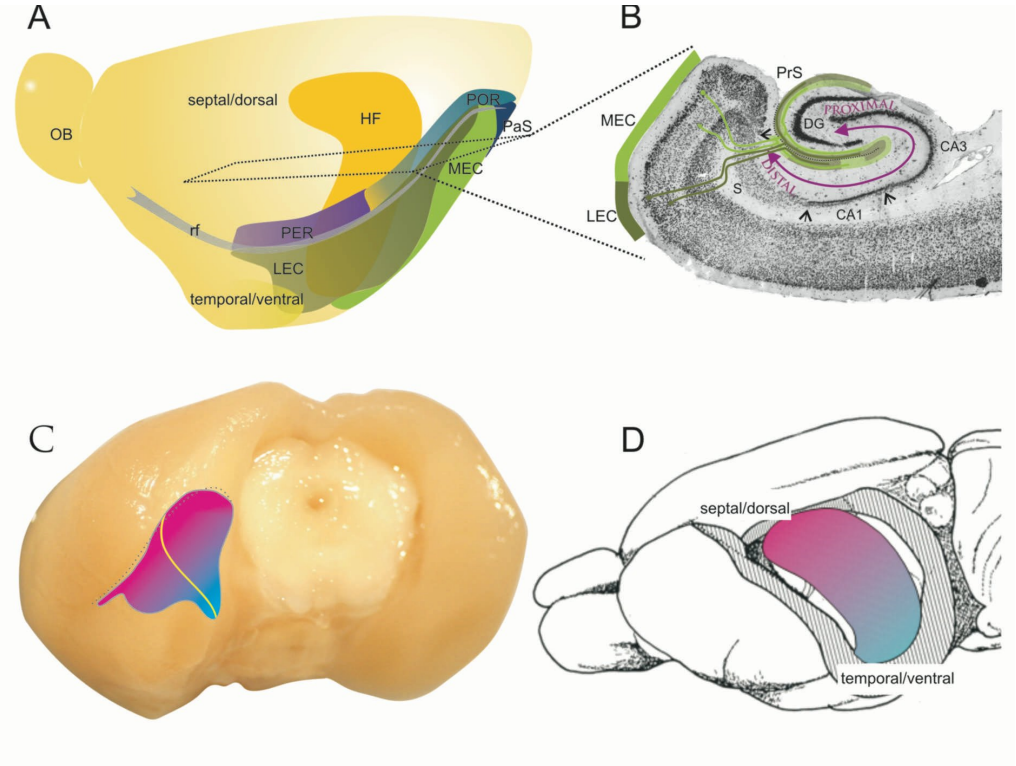


Figure 3.2. Anatomy of the entorhinal cortex (EC) in a rat brain, showing its main connectivity to the hippocampal formation (HF), the parasubiculum (PaS) and the presubiculum (PrS). The HF consists of dentate gyrus (DG), subiculum (S), and Cornu ammonis fields 1 to 3 (CA1, CA2 and CA3 respectively). The EC is divided in lateral entorhinal cortex (LEC) and medial entorhinal cortex (MEC). Other abbreviations: OB, olfactory bulb; POR, postrhinal cortex; rf, rhinal fissure; PER, perirhinal cortex. (A) Schematic lateral view of the left hemisphere of the rat brain. (B) Horizontal section taken from the dotted square in (A). Here the connections of the MEC and LEC with different parts of the HF are illustrated. (C) Position of the EC in the left hemisphere of the rat brain from a posterolateral view. Pink to blue shadow shows the dorsal to ventral dimension of the LEC and MEC. (D) Rat brain showing the HF of the right hemisphere. The colour-code shows the areas of the HF receiving input form dorsal (pink) and ventral (blue) EC. Figure modified from [25].

encode the animal's head direction in the horizontal plane, independent of location. Boundary cells fire whenever a boundary is at a particular distance and direction from the current location of the animal, independent of head direction. Some neurons in the lateral EC that typically fire in response to non-spatial cues, such as odour, can encode the relative distance and direction to the current or previous location of specific objects within an environment.

The grid pattern formed by grid cells firing locations (*firing fields*) can be described by three parameters: the *spacing* between the firing fields of a grid cell, the tilt of the grid relative to a reference axis (*orientation*) and the displacement in the x and y directions relative to an external reference point (*spatial phase*). The firing fields of neighbouring cells in MEC II have similar spacing and orientation, and different spatial phase [22]. When grid cells were discovered it was noticed that the size and spacing between these firing fields increase from dorsal to ventral positions in the MEC [22] (as illustrated in Fig. 3.3).

After the discovery of grid cells a number of computational models for grid-like firing patterns were proposed. Those models fall into two classes: oscillatory interference models and CAN models [17]. The first class uses interference patterns generated by multiple membrane-potential oscillations to explain grid formation [28] (see §2.4). The second class uses activity in local networks with specific connectivity to generate the grid pattern and its spacing [29, 30] (see §2.5.2). Recently these models have evolved due to new experimental findings and a new class of "self-organised" models has been proposed [31–33], whereby grid cells are formed by a self-organised learning process that borrows elements from both former classes.

The majority of grid cell models require the rat's running velocity and direction as an input for grid firing fields formation. However, the models proposed by Kropff and Treves [34], and Dordek *et al.* [35] demonstrated the emergence of grid cells without rat's velocity input, instead they used place cell inputs. These models are based on experimental evidence showing that grid cell activity disappears when place cells are inactivated [36], and that place cells develop earlier than grid cells in rat pups [37]. Interestingly, the model proposed by

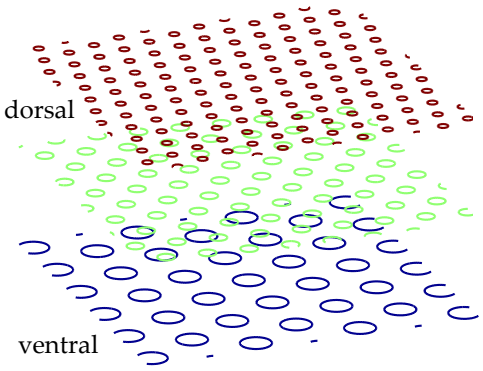


Figure 3.3. Scale, orientation, and phase of grid cell firing in Medial Entorhinal Cortex (MEC). Different colours represent different positions in the dorso-ventral axis of the MEC. The width of the circles shows the changes in scale and spacing of their firing fields along the dorso-ventral axis.

Dordek *et al.* exhibits a dependence between the place cell firing field size and the resulting grid spacing. This dependence is supported by experimental data exhibiting that the size of the place cell firing fields increases from dorsal to ventral positions of CA3 [38] (on similar fashion than the grid cell firing fields). The model in [35] is a single layer neural network with place-like input cells connected to grid cell outputs by feedforward weights that are learned via a Hebbian rule.

Conceptually, computational models of grid cells are useful to explain the formation of grid-like firing patterns, but in most cases they lack experimental evidence [19]. Moreover, some of these models fail to include the known inhibitory recurrent connectivity of the MEC II, or to explain the differences in the scale of firing fields of grid cells recorded at different locations of the dorsoventral axis of MEC [32]. For example, CAN models, such as those described in [26, 29, 39], are often criticised due the lack of evidence that patterns (or attractor states) are formed in the brain tissue. Additionally, experimental evidence shows that nearby grid cells in the tissue differ in their response features[2]; although nearby cells have the same spacing and orientation, they have different spatial phase [22]. This challenges the conception of a CAN model

2 Here the features of grid cell response are the spacing between the firing fields, their orientation and spatial phase.

because in these models neighboring nodes, representing neurons, have similar values in all their features. However, Yoon *et al.* [40] provided evidence that grid cells can be represented as CAN. In order to show this phenomena, Yoon *et al.* [40] analysed data from simultaneously recorded grid cells of a rat exploring an environment. First, they described the periodic tiling of the grid firing fields (emerging from different grid cells) with six parameters related to the angles (θ and ψ), length (λ_1 and λ_2) and location of their primary lattice vectors[3]. Figure 3.4 describes the first four parameters. The other two parameters are obtained by measuring the displacement of the central activity peak from the centre of the environment. Notably, cells recorded with the same or nearby tetrodes had similar values in four of these parameters. The parameters with different responses were those related to the location of the firing field centres, i.e., the spatial phase. Then, Yoon *et al.* modified the environment and recorded the activity of the same neurons. Although the individual cell responses changed, the grid parameter ratios and relative spatial phases between simultaneously recorded neurons remained constant, showing that relationships between cell pairs were conserved despite deformations of single neuron response. These experiments demonstrate that the recurrently connected network of grid cells is stable to deformations in the environment (that induce a change in a single cell activity). Importantly, because the cell-cell relationship is conserved when the single cell response changes it is hypothesised that the connectivity between cells plays a major role in the stability of grid cell response: without the connections between cells, the changes of single cell activity, induced by deformations to the environment, would destabilise the network and the firing fields of the deformed environments would not resemble the hexagonal firing fields obtained from grid cells. Interestingly, in CAN models the coupling between neurons plays a major role in the stability of the network [18, 19, 40], therefore the experimental evidence from Yoon *et al.* supports the hypothesis that the brain computes using continuous attractors. In CAN models of grid

3 The primary lattice vectors are related to the vertices of the hexagonal patterns formed at the locations in the rat's trajectory where the activity of a single grid cell is higher.

cells, the nodes are arranged in a 2D sheet with periodic boundary conditions, and the value of the features of grid cell response can be obtained from the position of the nodes. Note that the patterns generated in the 2D sheet of grid cells (nodes) correspond to states of activity in the network. These patterns are hexagonal because they represent the features of grid cells. Unlike the HD model described in §2.5.2, neighboring nodes in the model have close physical proximity in the brain.

In many of the continuous attractor models of grid cells, such as the models proposed by Couey *et al.* [26], and Burak and Fiete [39], a hexagonal activity pattern is formed in a 2D sheet of neurons (Fig. 3.5 Top). This hexagonal pattern is generated by the coupling between nodes, and its hexagonal shape represents the features of grid cell response. Once the pattern is formed, it is translated in response to velocity and head-direction signals that arise as the rat navigates an environment. This induces a single cell response, such as the one shown in Figure 3.1. The resulting firing fields can be obtained by recording the firing events (or cell activity) of a single cell in a 2D neural sheet at the rat's position in the environment. Figure 3.5 gives a schematic depiction of these models.

In this chapter it is assumed that the size and spacing of the grid cell firing fields are proportional to the size and spacing of the patterns generated in the 2D neural sheet. However, the computational models in [26, 39] only relate

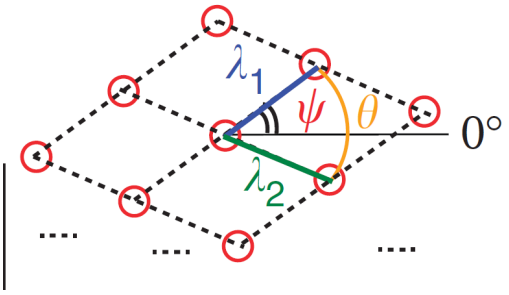


Figure 3.4. Spatial grid parameters used by Yoon *et al.* [40]. Here (λ_1, λ_2) describe the two primary axis lengths. Note that θ and ψ correspond to two different angles. Red dots correspond to the location of activity peaks recorded during an experiment. Figure taken from [40].

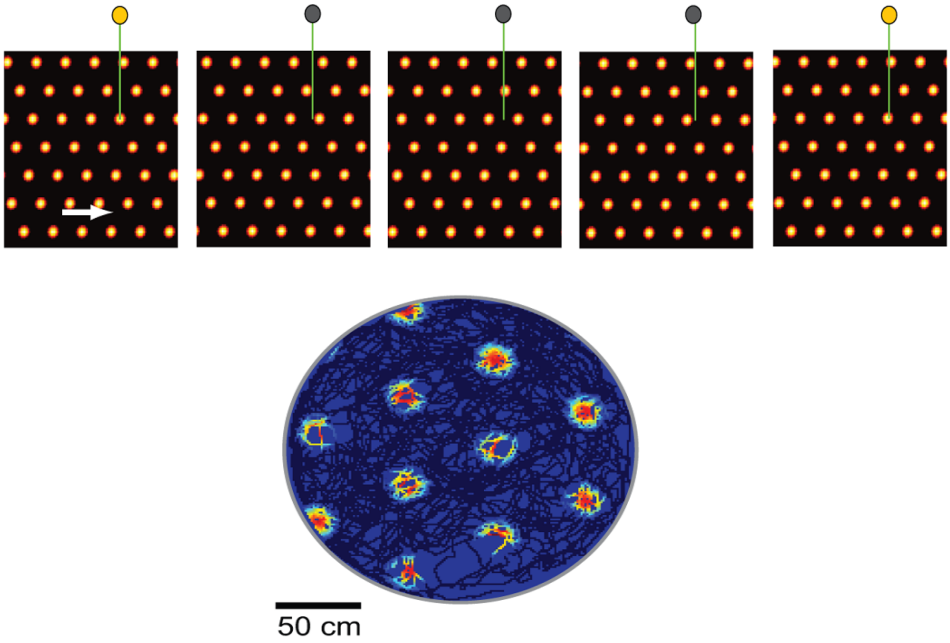


Figure 3.5. Continuous attractor model for grid cell activity. Top: Snapshots of the population activity at different times when the 2D network is driven by a constant velocity input in the rightward direction. The activity of a single neuron is recorded by an electrode (green line) and it is shown by the circle above the electrode. Here gray represents inactivity whereas yellow means that the cell is active. Bottom: Response avarege firing rate of the recorded cell, as a function of the simulated rat's position within the enclosure. Colour-coded red for high average firing rate and blue for low. Figure adapted from [39].

the change of size and spacing of the grid cell firing fields with the parameters that control the connectivity between cells. For example, simulations done by Couey *et al.* [26] exhibit larger firing fields when increasing the ratio of inhibition. On the other hand, Beed *et al.* [41] showed that there are more inhibitory inputs onto dorsal stellate cells in MEC II than ventral stellate cells using a combination of electrophysiological and optical approaches. Thus, the mechanism proposed by Couey *et al.* contradicts the experimental evidence given by Beed *et al.* because there is a strongly decreasing gradient of inhibition along the MEC dorsoventral axis while the size and spacing of the grid cell firing fields increase from dorsal to ventral positions along the MEC. Therefore the aim of this chapter is to find biological mechanisms and parameters not related to neural connectivity that control the size and spacing of grid cell firing fields.

3.2 A SIMPLE MODEL OF GRID CELL ACTIVITY

In this section a continuum version of the attractor model proposed first by Couey *et al.* [26] is studied, which can be written as

$$\tau \frac{\partial s(\mathbf{x},t)}{\partial t} + s(\mathbf{x},t) = gf\left(\int_{\mathbb{R}^2} W\left(\mathbf{x}-\mathbf{x}',l\right) s(\mathbf{x}',t)\mathrm{d}\mathbf{x}' + I + I_{vel}(\mathbf{x},t)\right). \tag{3.1}$$

Here $s = s(\mathbf{x},t)$ is the measure of neural activity as a function of the neuron location on a two dimensional plane, $\mathbf{x} = (x,y) \in \mathbb{R}^2$, at time t; f is a threshold-linear function defined by

$$f(z) = \begin{cases} z, & \text{if } z > 0 \\ 0, & \text{if } z \leq 0 \end{cases}, \tag{3.2}$$

g is constant gain; and τ is the neural time constant. I is a constant external input that could be generated by direct or indirect connections from the hippocampus. The model receives head-directional input tuned to its preferred direction through the term $I_{vel}(\mathbf{x},t) = \alpha v(t)\cos\left(\varphi(t) - \theta(\mathbf{x})\right)$, where α is the velocity modulation; $v(t)$ is the animal velocity at time t; and $\varphi(t)$ is the animal head direction at time t. Here, $\theta(\mathbf{x})$ represents the preferred head direction

of each neuron, so that there is an increase in firing rate when the animal is oriented this way.

Motivated by experimental observations, Couey *et al.* proposed that neurons are connected to each other via a pattern that follows a simple all-or-none unstructured inhibitory connectivity defined by

$$W(\mathbf{x},\theta) = W_0 H\left(R - \sqrt{(x - l\cos(\theta))^2 + (y - l\sin(\theta))^2}\right), \tag{3.3}$$

where H is the Heaviside function given by (2.23). Thus each neuron only receives inhibition of constant magnitude (i.e. $W_0 < 0$) within a radius of connectivity, R. Note that in this CAN model, the distance between neurons are related to the features of grid cells response (the spacing and size of firing fields and their relative phases) and not to the physical locations in the neural tissue. Therefore, the size of R represents the number of inhibitory inputs to the node located at $\mathbf{x}$ in the 2D sheet. Furthermore, the connectivity has an offset, l, in the network space according to the preferred direction of each cell, $\theta = \theta(\mathbf{x})$. The combination of the head-directional input and the offset enables the grid activity on the network to translate with the animal's movement, as illustrated in Figure 3.5. Couey *et al.* hypothesised that this type of connectivity forms stable hexagonal patterns when there is no head-directional input or offset (i.e., when $\alpha = l = 0$) because the neurons only inhibit other neurons nearby. This inhibition may be exceeded by input (I); however, the remaining activity inhibits nearby firing via (3.3) so that areas of activity settle in positions where the distance between them is maximised, namely the vertices of a hexagonal grid. However, the next section shows that this type of connectivity, with $\alpha = l = 0$, can also form stripe patterns.

Here, it is assumed that the rat is stationary (setting $\alpha = 0$) in order to study the scale of the emergent patterns. Because there is no head-direction input, the patterns do not translate and the model can be further simplified by assuming $l = 0$. Thereby, the continuous attractor model proposed by Couey

et al. is now simplified to an attractor model. This simplified model can be written as:

$$\tau \frac{\partial s(\mathbf{x},t)}{\partial t} + s(\mathbf{x},t) = gf\left(\int_{\mathbb{R}^2} W\left(\mathbf{x}-\mathbf{x}'\right) s(\mathbf{x}',t)\mathrm{d}\mathbf{x}' + I\right), \tag{3.4}$$

where

$$W(\mathbf{x}) = W_0 H\left(R - \sqrt{x^2+y^2}\right) = W_0 H\left(R - |\mathbf{x}|\right). \tag{3.5}$$

Considering that current pattern formation theory assumes the firing rate function to be differentiable everywhere (see Appendix B and [8,42] for details), it is replaced by the sigmoid function suggested in [43]

$$f(x) = \mu \ln^{\beta}\left(1 + \mathrm{e}^{b(x+c)}\right). \tag{3.6}$$

Note that this function behaves like x for large $x > 0$ and for large $x < 0$ it behaves like $\mathrm{e}^{-|x|}$, making a good smooth approximation to the threshold-linear function (Fig. 3.6a). Figure 3.6b shows that this choice of f forms hexagonal patterns in direct numerical simulations of the attractor model.

3.2.1 *Linear Stability Analysis*

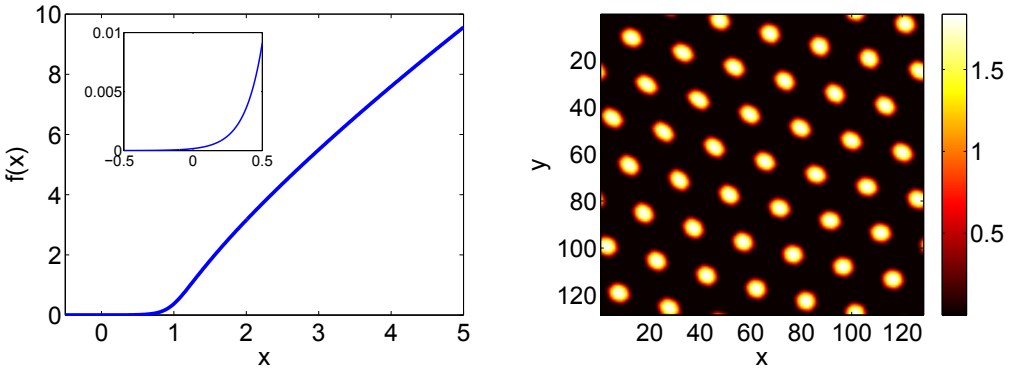
In order to investigate the conditions under which patterns can arise from this model a linear stability analysis is performed by introducing small spatially heterogeneous perturbations around a spatially homogeneous steady state. The linearised system solution is a good predictor of the emergent patterns close to bifurcation. Patterns are formed if such system is unstable to perturbations that grow exponentially in time. This growth is not indefinite, it is saturated in the long-term by the nonlinear terms of the full model.

First, let $s(\mathbf{x},t) = \overline{s}$ be the spatially homogeneous steady state. Substitution into (3.4) gives

$$\overline{s} = gf\left(\overline{s}\int_{\mathbb{R}^2} W(\mathbf{x}-\mathbf{x}')\mathrm{d}\mathbf{x}' + I\right) = gf\left(\overline{s}\,\overline{W} + I\right), \tag{3.7}$$

where

$$\overline{W} \;=\; W_0 \int_{\mathbb{R}^2} H\left(R - |\mathbf{z}|\right) \mathrm{d}\mathbf{z} = W_0 \int_0^{2\pi}\int_0^R r\mathrm{d}r\mathrm{d}\phi = \pi W_0 R^2 < 0, \tag{3.8}$$



(a) (b)

Figure 3.6. (a) Sigmoid firing rate function in (3.6). Inset: Zoom of the function near the origin. Parameters: $\mu = 0.5$, $\beta = 0.8$, $b = 10$, $c = -1$. (b) Simulation snapshot at $t = 500$ms of a simple model of grid cell activity $s(\mathbf{x}, t)$ in (3.4) using the sigmoid rate function in (a). Parameters: $R = 15$, $W_0 = -0.02$, $\mu = 0.5$, $\tau = 10$, $g = 1$, $I = 3$. Simulation details are given in Appendix C.

and $r = |\mathbf{x}|$. Figure 3.7 shows a plot of the steady state. Note that the steady state always exists due the negative slope of the right hand side of (3.7) induced by $\overline{W}$.

Then, linearisation around the steady state, by letting $s(\mathbf{x}, t) \to \overline{s} + s(\mathbf{x}, t)$, $s(\mathbf{x}, t) \ll 1$ in (3.4), and taking the Taylor expansion of f about $\overline{W}\overline{s} + I$ gives (after dropping the quadratic terms)

$$\tau \frac{\partial s(\mathbf{x}, t)}{\partial t} + s(\mathbf{x}, t) = g\gamma \left(\int_{\mathbb{R}^2} W\left(\mathbf{x} - \mathbf{x}'\right) s(\mathbf{x}', t) \mathrm{d}\mathbf{x}' \right), \tag{3.9}$$

where $\gamma = f'(\overline{W}\overline{s} + I)$. This linearised system has solutions of the form

$$s(\mathbf{x}, t) = A \mathrm{e}^{i\mathbf{k}\cdot\mathbf{x}} \mathrm{e}^{\lambda t}, \qquad \mathbf{k} = (k_1, k_2) \in \mathbb{R}^2, \qquad \lambda \in \mathbb{C}, \tag{3.10}$$

where λ is the *growth-rate* and $k = |\mathbf{k}|$ is the *wave-number*. Note that the perturbation is spatially periodic with wavelength $2\pi/k$. This perturbation either grows or decays exponentially in time depending on $\nu = \mathrm{Re}(\lambda)$ with the emergent temporal frequency of oscillations given by $\omega = \mathrm{Im}(\lambda)$. If $\max(\mathrm{Re}(\lambda)) < 0$ for all $k \in \mathbb{R}, k \neq 0$, then the steady state is linearly stable because the perturbation tends to zero in the limit $t \to \infty$. An instability occurs when for the first time there are values of k for which the real part of λ is non-negative; these

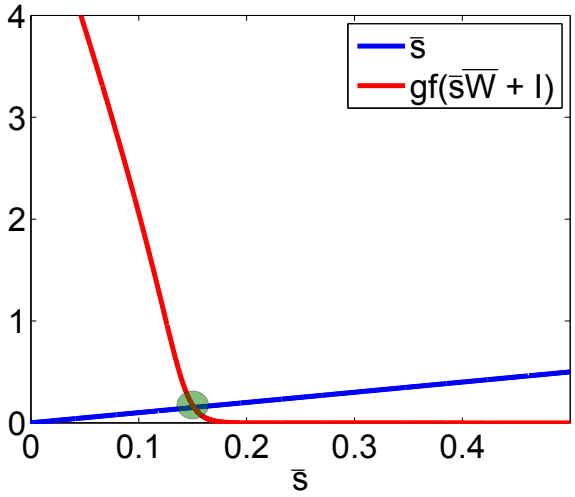


Figure 3.7. Steady State of a simple model of grid cell activity (3.8). Here the green circle represents the intersection between $\bar{s}$ (blue line) and $gf\left(\bar{s}\overline{W}+I\right)$ (red line). Parameters as in Figure 3.6

values are termed 'critical' and denoted k_c, $\lambda_c = \nu_c + i\omega_c$. A *Turing bifurcation point* is defined as the smallest value of some parameter for which there exists some non-zero critical wave-number k_c satisfying $\nu_c = 0$. It is said to be *static* if $\omega_c = 0$ and *dynamic* if $\omega_c \neq 0$. The dynamic instability is often referred to as a *Turing-Hopf bifurcation* and generates a global pattern with wave-number k_c, which moves coherently with a speed $c = \omega_c/k_c$, i.e. as a periodic travelling wave train. If the maximum of the dispersion curve[4] is at $k_c = 0$ then the mode that is first excited is another spatially uniform state. If $\omega_c = 0$ then the system evolves to another homogeneous state since no pattern exists. This type of instability is called *bulk instability*. If $\omega_c \neq 0$, a coherent network oscillation emerges with frequency ω_c. Substituting (3.10) in (3.9) gives the following so-called *dispersion relation* between λ and k

$$\lambda(k) = \frac{1}{\tau}\left(g\gamma\widetilde{W}(k;R,W_0) - 1\right). \tag{3.11}$$

Here

$$\begin{aligned}\widetilde{W}(k;R,W_0) &= W_0\int_0^R\int_0^{2\pi} \mathrm{e}^{irk\cos(\phi)} r\mathrm{d}\phi\mathrm{d}r \\ &= W_0 2\pi\int_0^R J_0(rk)r\mathrm{d}r = \frac{W_0 2\pi R}{k}J_1(Rk),\end{aligned} \tag{3.12}$$

4 Here the dispersion curve is the plot of $\mathrm{Re}\,(\lambda)$ against k.

is the 2D Fourier transform of W. Here J_l, $l \in \{0, 1\}$, denotes the Bessel function of the first kind given by

$$J_n(z) = \frac{1}{2\pi i^n} \int_0^{2\pi} \mathrm{e}^{iz\cos(\phi)} \mathrm{e}^{in\phi} \mathrm{d}\phi, \tag{3.13}$$

and the fact that W is radially symmetric has been used. Note that $\widetilde{W}(k; R, W_0) \in \mathbb{R}$, $\forall k$, therefore $\lambda(k) \in \mathbb{R}$, $\forall k$. A plot of $\widetilde{W}(k; R, W_0)$ is shown in Figure 3.8.

The homogeneous steady state is stable if $\lambda(k) < 0$ $\forall k$. Figure 3.8 shows that $\widetilde{W}(k)$ has a positive maximum at $\pm k_c$. Therefore there is a bifurcation point $\gamma = \gamma_c = 1/(g\widetilde{W}(k_c))$ such that $\lambda(k) < 0$ for all $k \neq k_c$ with $\lambda(k_c) = 0$. At this point the system goes from stable with $\gamma < \gamma_c$ to unstable where $\gamma > \gamma_c$ and there exists a range of values of $k \in (k_{c_1}, k_{c_2})$ for which $\lambda(k) > 0$, as shown in Figure 3.9.

At the bifurcation point the system becomes unstable due to the excitation of the pattern $\mathrm{e}^{i\mathbf{k}_c \cdot \mathbf{x}}$, and beyond this bifurcation point there is a set of growing patterns that are ultimately determined by the full nonlinear model. Since $\lambda(k) \in \mathbb{R}$ $\forall k$, the bifurcation pattern is static and spatially periodic with *wavelength* $2\pi/k_c$ (Figs. 3.6b, 3.10). Furthermore, it can be predicted if the emergent patterns will form stripes or spots near the bifurcation point (see Appendix B). For example, Figure 3.10 shows the different types of patterns that can arise depending on the parameters of the firing rate.

If the gain, g, or magnitude of inhibition, W_0, are varied in the proposed model, the critical wave-number will remain the same. In fact the only way to alter k_c is by changing the connectivity radius because this is the only parameter that moves the location of $\max_k \widetilde{W}(k)$. It can be seen in Figure 3.11a that when R (which relates to the number of inhibitory inputs) is increased k_c decreases, and therefore the wavelength of the patterns increments, making the grid spacing and size bigger. Furthermore, in Figure 3.11b it is shown that the relationship between R and $1/k_c$ is linear. However, experimental evidence shows that the inhibitory inputs in stellate cells of the MEC decreases from dorsal to ventral parts [41] and that the size and spacing of grid cell firing fields increase along the dorso-ventral axes [22], exhibiting an opposite trend

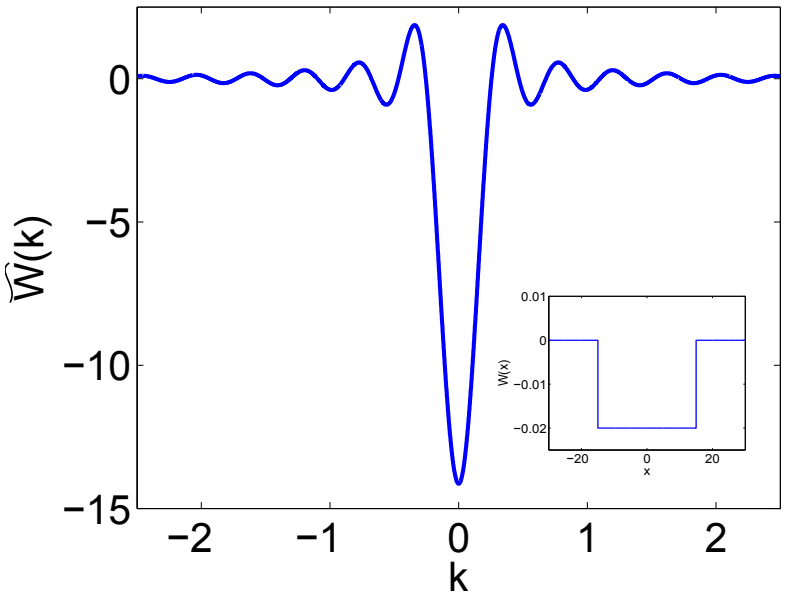


Figure 3.8. The 2D Fourier transform ($\widetilde{W}(k; R, W_0)$ defined in (3.12)) of the inverse Top Hat function plotted in the inset ($W(x; R, W_0)$ defined in (3.5)). Parameters: $R = 15$, $W_0 = -0.02$.

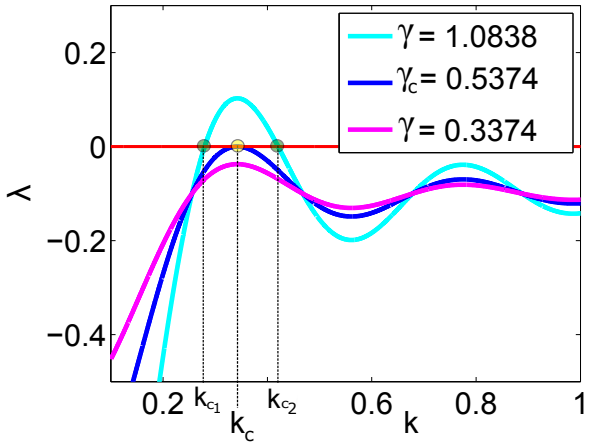
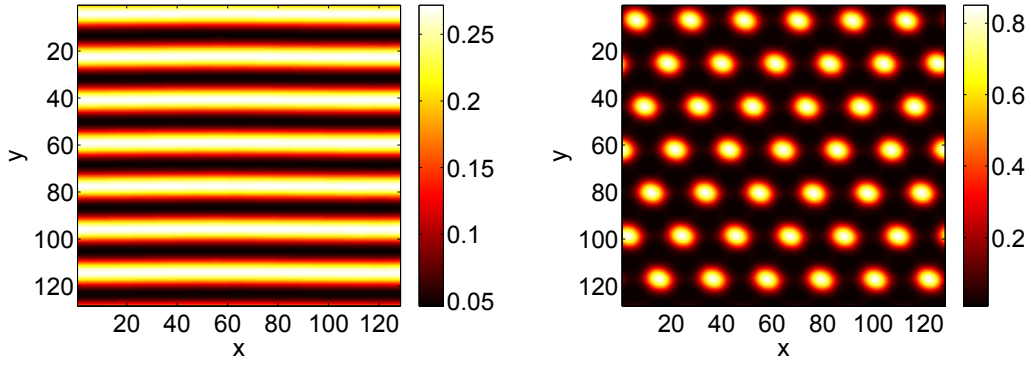


Figure 3.9. The eigenvalue λ of (3.11) as a function of k for different values of γ (corresponding to a simple model for grid cell dynamics (3.4)). The yellow circle represents the bifurcation point ($\lambda(k_c) = 0$) whereas the green ones indicates the set of $k \in (k_{c_1}, k_{c_2})$ where the system is unstable ($\lambda(k) > 0$). Parameters: $R = 15$, $W_0 = -0.02$, $\tau = 10$, $g = 1$, $I = 3$. For the firing rate function: $\mu = 0.5$, $\beta = 0.8$ and for the cyan, blue and magenta lines $b = 10, 4.9340, 3.0886$, $c = -1, -1.1417, -1.3089$ respectively.



(a) (b)

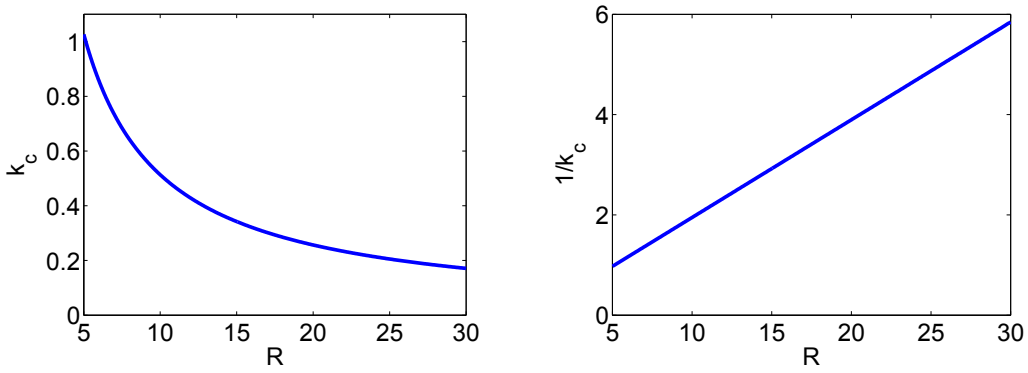
Figure 3.10. Resulting patterns in a simple model of grid cell activity ($s(\mathbf{x}, t)$ in (3.4)) after $t = 1500$ms. Parameters: $\tau = 10$, $R = 15$, $W_0 = -0.02$, $g = 1$, $I = 3$, $\beta = 0.8$, $\gamma_c = 0.5347$, $k_c = 0.3423$. (a) Striped pattern using $b = 10$, $\mu = 0.0891$ and $c = -0.6839$. (b) Spotted pattern using $b = 4.9340$, $c = -1.1417$ and $\mu = 0.5$. Details for the simulation are in Appendix C, and the details for pattern selection are found in Appendix B.

from that described by the proposed model. Thus changing R is an unrealistic biological mechanism to control the size of patterns.

In the following sections other possible mechanisms for grid cell dynamics are proposed. First, in §3.3 adaptation is added to this simple model, and then second order synapses are implemented in §3.3.2. Then, §3.4 investigates if the delays arising from neural interactions can account for the difference of pattern scales in the grid cell firing fields. In §3.5 resonances are included, and in §3.6 dendrites (and various types of axo-dendritic connections) are modelled. Finally, in §3.7 resonances are included to the axo-dendritic connections in §3.6 in order to investigate if local tissue parameters unrelated to the network connectivity can give rise to long wavelength patterns.

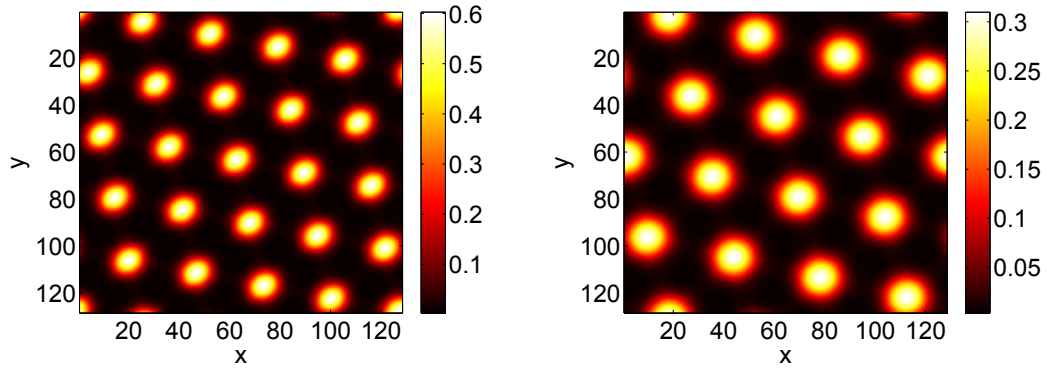
3.3 MODEL WITH ADAPTATION

Some neurons have an intrinsic negative feedback that lowers the firing rate of a neuron called *spike-frequency adaptation*. Curtu and Ermentrout showed



(a) (b)

Figure 3.11. The critical wave-number k_c ($\lambda(k_c) = 0$ in (3.11)) (a) and its reciprocal (b) as a function of the connectivity radius R for a simple model of grid cell activity (3.4). Here the new steady state (3.7) is calculated for each radius and the firing rate function parameters b and c (3.6) are changed accordingly in order to get the critical value of γ_c. Other parameters: $W_0 = -0.02$, $\tau = 10$, $g = 1$, $I = 3$. For the firing rate function: $\mu = 0.5$, $\beta = 0.8$.



(a) (b)

Figure 3.12. Resulting patterns in a simple model for grid cell activity ($s(\mathbf{x}, t)$ in (3.4)) after $t = 1500$ms. Parameters: $\tau = 10$, $W_0 = -0.02$, $g = 1$, $I = 3$, $\beta = 0.8$. (a) $R = 20$, $\mu = 0.5$, $b = 4.5213$, $c = -1.2562$, $k_c = 0.2568$ and $\gamma_c = 0.3008$ (b) $R = 25$, $\mu = 0.1746$, $b = 4.7134$, $c = -1$, $k_c = 0.2054$ and $\gamma_c = 0.1925$. Note that the pattern wavelength increases due to the decrease in k_c when increasing R (see Fig. 3.11). Details for the simulation and pattern selection are in Appendix C and Appendix B respectively.

in [44] that different patterns can emerge from a network of excitatory and inhibitory cells with adaptation depending on the strength and time constant of adaptation. Hence, an adaptation term is added to the simplified model (3.4). It is investigated if patterns can be formed in the inhibitory network (3.5) and if so whether their scale can be related to adaptation parameters. The new system is given by:

$$\tau \frac{\partial s(\mathbf{x},t)}{\partial t} + s(\mathbf{x},t) = gf\left(\int_{\mathbb{R}^2} W\left(\mathbf{x}-\mathbf{x}'\right) s\left(\mathbf{x}',t\right) \mathrm{d}\mathbf{x}' + I - ha(\mathbf{x},t)\right), \tag{3.14a}$$

$$\tau_a \frac{\partial a(\mathbf{x},t)}{\partial t} + a(\mathbf{x},t) = s(\mathbf{x},t), \tag{3.14b}$$

where $a(\mathbf{x},t)$ represents the neural adaptation and $h, \tau_a > 0$ represent the strength and time constant of adaptation, respectively. Typically adaptation is thought of as a slow local negative feedback, hence $\tau_a > \tau$.

3.3.1 *Linear Stability Analysis*

Let $(s(\mathbf{x},t), a(\mathbf{x},t)) = (\overline{s}, \overline{a})$ be the spatially homogeneous steady state, then substituting in (3.14) gives

$$\overline{s} = gf\left(\overline{W}\overline{s} + I - h\overline{a}\right), \qquad \overline{a} = \overline{s}, \tag{3.15}$$

where $\overline{W}$ is defined in (3.8). Linearisation around the steady state is performed by letting $(s(\mathbf{x},t), a(\mathbf{x},t)) \to (\overline{s}, \overline{a}) + (s(\mathbf{x},t), a(\mathbf{x},t))$. Substituting into (3.15) and seeking solutions of the form

$$(s(\mathbf{x},t), a(\mathbf{x},t)) = (s_0, a_0)\, \mathrm{e}^{i\mathbf{k}\cdot\mathbf{x}} \mathrm{e}^{\lambda t}, \qquad \mathbf{k} = (k_1, k_2) \in \mathbb{R}^2, \tag{3.16}$$

gives (after Taylor expanding)

$$\begin{aligned} \lambda\tau s_0 + s_0 &= s_0 g\gamma\left(\widetilde{W}(k; R, W_0) - ha_0\right), \\ \lambda\tau_a a_0 + a_0 &= s_0. \end{aligned}$$

Here $\widetilde{W}(k; R, W_0)$ is given by (3.12) and $\gamma = f'\left(\overline{W}\overline{s} + I - h\overline{a}\right)$. The above system can be rewritten in a matrix form

$$\begin{bmatrix} \lambda\tau + 1 - g\gamma\widetilde{W} & g\gamma h \\ -1 & \lambda\tau_a + 1 \end{bmatrix} \begin{bmatrix} s_0 \\ a_0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \tag{3.17}$$

Requiring non-trivial solutions, the following characteristic polynomial is obtained:

$$\tau\tau_a\lambda^2 + \alpha_0\lambda + \alpha_1 = 0, \tag{3.18}$$

where $\alpha_0 = \tau_a\left(1 - g\gamma\widetilde{W}\right) + \tau$ and $\alpha_1 = 1 - g\gamma\widetilde{W} + g\gamma h$. The values of λ in (3.18) can be either real or complex. Therefore the linear system will destabilise when one or both eigenvalues become positive. Writing $\lambda = \nu + i\omega$, $\nu, \omega \in \mathbb{R}$, and substituting into (3.18) gives the following system

$$\tau\tau_a\nu^2 - \tau\tau_a\omega^2 + \alpha_0\nu + \alpha_1 = 0, \tag{3.19a}$$

$$2\tau\tau_a\nu\omega + \alpha_0\omega = 0. \tag{3.19b}$$

Note that the steady state of the system, given by (3.15), depends on the choice of g, I, the connectivity parameters (R and W_0) and the strength of adaptation h. Hereinafter a parameter not related to the steady state is chosen to be the bifurcation parameter, in order to not interfere with the existence of the steady state. Here, τ_a is the bifurcation parameter. At the bifurcation point $\nu = 0$, $k = k_c$ and $\tau_a = \tau_{a_c}$ in (3.19), which gives

$$\tau_{a_c}\left(1 - g\gamma\widetilde{W}(k_c)\right) + \tau = 0, \qquad \omega_c^2 = \frac{1 - g\gamma\widetilde{W}(k_c) + g\gamma h}{\tau_{a_c}\tau} > 0. \tag{3.20}$$

As in the previous model (3.4) $\widetilde{W}(k_c) = \max_k \widetilde{W}(k)$. It can be seen in Figure 3.13 that when $\tau_a < \tau_{a_c}$ (magenta line), $\nu(k) < 0$ $\forall k$, therefore the system is stable. Whereas if $\tau_a > \tau_{a_c}$ (cyan line) then there is a range of k for which the system is unstable. The curves in Figure 3.13 are described by (after solving (3.19a) and (3.19b))

$$\nu^2 + \omega^2 + \frac{2\nu}{\tau_a} = \frac{g\gamma h}{\tau\tau_a} - \frac{1}{\tau_a^2}. \tag{3.21}$$

In contrast to the model without adaptation (3.4), here the bifurcation is dynamic and has an emergent temporal frequency of $\omega = \sqrt{\left(-\tau/\tau_{a_c} + g\gamma h\right)/\left(\tau\tau_{a_c}\right)}$.

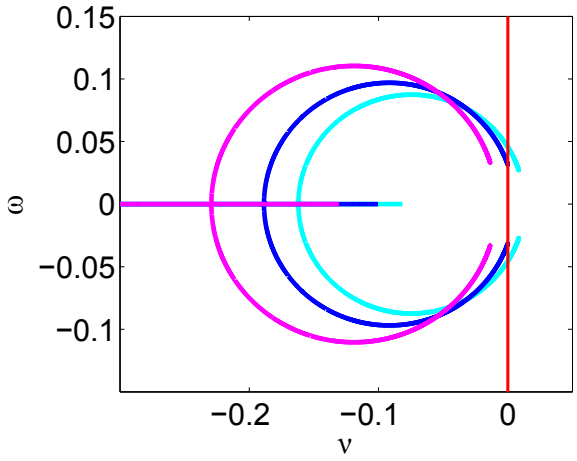


Figure 3.13. Spectral picture given by (3.21) of the model with adaptation (3.14) for different values of the adaptation time constant, $\tau_a = 8.4095$ (magenta), 10.9095 (blue) and 13.4095 (cyan). Parameters: $R = 15$, $W_0 = -0.02$, $\tau = 10$, $g = 1$, $I = 3$, $h = 1$; and for the firing rate function: $\beta = 0.8$, $\mu = 0.5$, $b = 10$ and $c = -1$.

Moreover, varying other parameters, (such as τ, data omitted), do not affect the critical wave-number since it only depends on R (via $\widetilde{W}(k)$ as in a simple model for grid cell activity in §3.2). Therefore adaptation fails to provide a mechanism to explain the change in size and spacing of grid cells.

3.3.2 *Model with Adaptation and Second Order Synapses*

Here, second order synapses are introduced into (3.14a), yielding the new system:

$$\left(1+\tau\frac{\partial}{\partial t}\right)^2 s(\mathbf{x},t) = gf\left(\int_{\mathbb{R}^2} W\left(\mathbf{x}-\mathbf{x}'\right) s\left(\mathbf{x}',t\right) \mathrm{d}\mathbf{x}' + I - ha(\mathbf{x},t)\right), \tag{3.22a}$$

$$\left(1+\tau_a\frac{\partial}{\partial t}\right) a(\mathbf{x},t) = s(\mathbf{x},t). \tag{3.22b}$$

As in §3.3.1 a linear stability analysis is performed and the non-trivial solutions of the linearised system are given by the dispersion relation $\varepsilon(\lambda, k) = 0$ with

$$\begin{aligned}\varepsilon(\lambda,k) &= \lambda^3\tau^2\tau_a + \lambda^2\left(2\tau\tau_a + \tau^2\right) + \lambda\left(2\tau + \tau_a\left(1 - g\gamma\widetilde{W}(k)\right)\right) \\ &\quad -g\gamma\widetilde{W}(k) + 1 + gh\gamma. \end{aligned} \tag{3.23}$$

The dispersion relation can have either 3 real or 1 real and 1 complex conjugate pair of eigenvalues; stability is lost when the real part of one of the eigenvalues becomes positive. Assuming $\lambda \in \mathbb{C}$, $\lambda = \nu + i\omega$, the dispersion relation in (3.23) can be rewritten as

$$\varepsilon(\nu + i\omega, k) \equiv \mathcal{G}(\nu, \omega, k) + i\mathcal{H}(\nu, \omega, k) = 0, \tag{3.24}$$

where

$$\begin{aligned}\mathcal{G}(\nu,\omega,k) &= \mathrm{Re}\left(\varepsilon(\nu + i\omega, k)\right) = \tau^2\tau_a\left(\nu^3 - 3\nu\omega^2\right) + \left(2\tau\tau_a + \tau^2\right)\left(\nu^2 - \omega^2\right) \\ &\quad + \left(2\tau + \tau_a\left(1 - g\gamma\widetilde{W}(k)\right)\right)\nu - g\gamma\widetilde{W}(k) + 1 + gh\gamma, \end{aligned} \tag{3.25a}$$

$$\begin{aligned}\mathcal{H}(\nu,\omega,k) &= \mathrm{Im}\left(\varepsilon(\nu + i\omega, k)\right) = \tau^2\tau_a\left(3\nu^2\omega - \omega^3\right) + \left(2\tau\tau_a + \tau^2\right)2\nu\omega \\ &\quad + \left(2\tau + \tau_a\left(1 - g\gamma\widetilde{W}(k)\right)\right)\omega. \end{aligned} \tag{3.25b}$$

Note that (3.25) is not, on its own, a system providing ω_c, k_c and the bifurcation parameter τ_{a_c}. A third equation can be obtained considering that at the bifurcation point $\mathrm{d}\nu/\mathrm{d}\omega = 0$. The implicit function theorem gives the following equivalent condition:

$$\frac{\mathrm{d}\mathcal{G}}{\mathrm{d}\omega}\frac{\mathrm{d}\mathcal{H}}{\mathrm{d}k} - \frac{\mathrm{d}\mathcal{H}}{\mathrm{d}\omega}\frac{\mathrm{d}\mathcal{G}}{\mathrm{d}k} = 0. \tag{3.26}$$

The system of equations (3.25) and (3.26) is numerically solved using MATLAB's `fsolve` function. The linearised system loses stability when $\tau_a > \tau_{a_c}$ and the complex conjugate pair of eigenvalues become positive as shown in Figure 3.14. Here k_c corresponds to the value of k that maximises $\widetilde{W}(k)$ and it remains the same when varying other model parameters not related to the connectivity.

When computing the solution to (3.25) and (3.26), care must be taken to ensure that the critical point obtained corresponds to $\max_k \widetilde{W}(k)$ rather than

to $\min_k \widetilde{W}(k)$. Including adaptation and second order synapses in the model results in a dynamical bifurcation but not a mechanism to change the size and spacing of the patterns.

3.4 MODEL WITH DELAYS

So far, the modifications to the simple model for grid cell activity (3.4) have failed to give a mechanism by which the size and spacing of the patterns can be controlled by parameters independent of the connectivity. Following the work of Roxin *et al.* [45], in this section a delay, D, is added to the simplified model (3.4) in order to investigate its effects. The interactions between neurons involve delays on the order of milliseconds due to the velocity of the propagation of action potential or processes within the neurons. The new model is given by:

$$\tau \frac{\partial s(\mathbf{x},t)}{\partial t} + s(\mathbf{x},t) = gf\left(\int_{\mathbb{R}^2} W\left(\mathbf{x}-\mathbf{x}'\right) s\left(\mathbf{x}', t-D\right) \mathrm{d}\mathbf{x}' + I\right). \tag{3.27}$$

Let $s(\mathbf{x},t) = \overline{s}$ be the spatially-homogeneous steady state, then substituting into (3.27) gives (3.7). Linearising about the steady state and looking for solutions of the form (3.10) gives the following dispersion relation

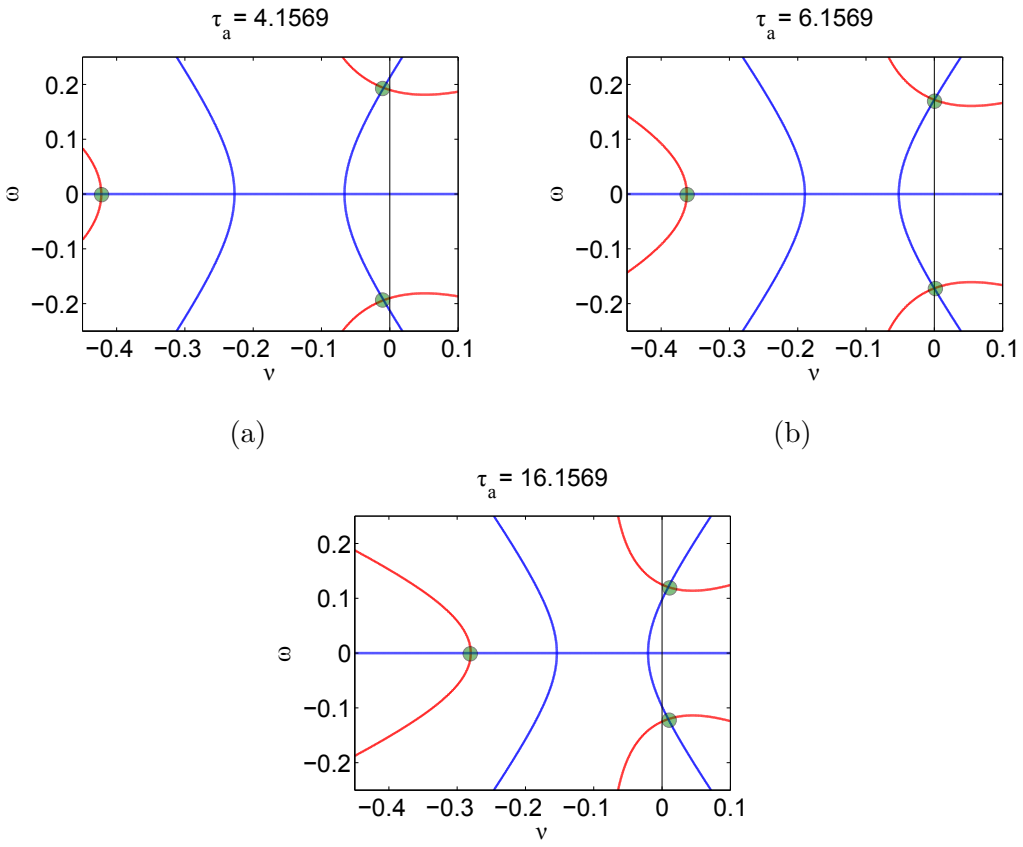
$$\varepsilon(\lambda,k) = \lambda\tau + 1 - g\gamma \mathrm{e}^{-\lambda D}\widetilde{W}(k) = 0, \tag{3.28}$$

where $\gamma = f'(\overline{W}\overline{s} + I)$. Assuming $\lambda \in \mathbb{C}$, with $\lambda = \nu + i\omega$, then the linearised system has a dynamic bifurcation when $\nu = 0$ and

$$\tau\omega = -\tan(\omega D_c). \tag{3.29}$$

Figure 3.15a shows the spectral picture at the bifurcation point with D_c given by (3.29). Although the chosen D_c and $\omega_c \neq 0$ in Figure 3.15a correspond to the solution of (3.29) (green circles at $\nu = 0$), the system has an eigenvalue that is already unstable (green circle at $\nu > 0$). Such an instability occurs when $\lambda \in \mathbb{R}$, $\lambda = 0$ and

$$1 = g\gamma_c \widetilde{W}(k_c). \tag{3.30}$$



(a) (b)

(c)

Figure 3.14. Spectral picture for the model with adaptation and second order synapse (3.22) for diffent values of τ_a. Red lines represent $\mathcal{G}(\nu, \omega, k) = 0$ (3.25a), blue lines $\mathcal{H}(\nu, \omega, k) = 0$ (3.25b), the green circles the eigenvalues. Parameters: $W_0 = -0.02$, $R = 15$, $\tau = 10$, $g = 1$, $I = 3$, $h = 10$. For the firing rate function: $\beta = 0.8$, $\mu = 0.5$, $b = 10$, $c = -1$. In (a) the system is stable, (b) is at bifurcation point and in (c) the complex conjugate eigenvalues are unstable.

Note that (3.30) is the same condition as the one for the simple model for grid cell activity (3.11). Hence, k_c is the value that maximises $\widetilde{W}(k)$ and its value remains the same when varying the other parameters.

In conclusion, the addition of delays in this model did not fulfil its purpose, namely to find a mechanism to vary the size and spacing of the patterns. This is because the system loses stability when the real eigenvalue becomes positive before the dynamic bifurcation from (3.29) occurs. The bifurcation point is the same as in the model described in §3.2 and, as discussed there, changing parameters other than the connectivity radius do not affect the critical wavenumber and therefore the pattern scale.

3.5 RESONANT MODEL

In vitro experimental evidence shows that stellate cells in dorsal Entorhinal Cortex exhibit higher temporal frequencies of subthreshold membrane potential oscillations compared to lower frequencies in cells from more ventral Entorhinal slices [46]. Hence, in this section a model combining elements of quasi-active membrane models (§2.2.2) that can account for subthreshold oscillations, with network dynamics (§2.5) is proposed. Suppose a neuron at position $\mathbf{x} = (x, y)$ has voltage $V(\mathbf{x}, t)$ and current $I(t)$ that evolve according to

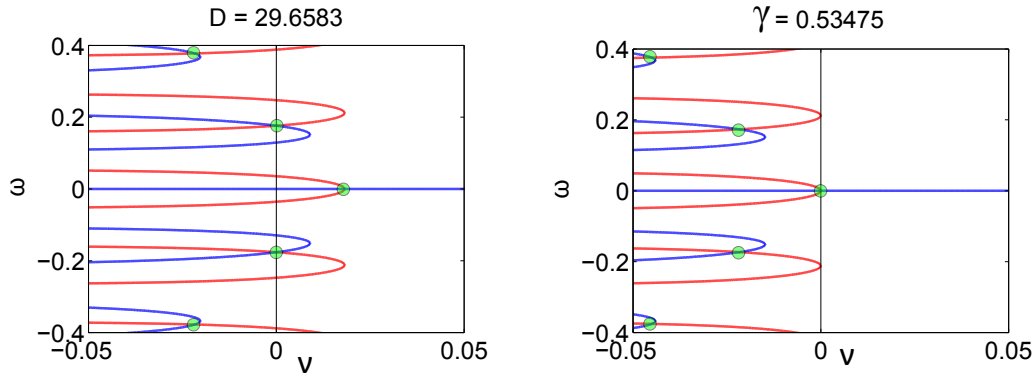
$$C\frac{\mathrm{d}V}{\mathrm{d}t} = -\frac{V}{R} + I_{syn} - I, \tag{3.31a}$$

$$L\frac{\mathrm{d}I}{\mathrm{d}t} = V - rI, \tag{3.31b}$$

where

$$I_{syn} = \int_{-\infty}^{t} \mathrm{d}t'\, \eta(t-t') \int_{\mathbb{R}^2} \mathrm{d}\mathbf{x}'\, W(\mathbf{x}-\mathbf{x}') f(V(\mathbf{x}', t')) \equiv \eta * W \otimes f(V). \tag{3.32}$$

Here I_{syn} represents synaptic activity (see §2.5.1) and η is the α-function given in (2.22). The temporal convolution is represented by $*$ whereas $\otimes$ describes spatial convolution.



(a) (b)

Figure 3.15. Spectral picture for the model with delays (3.27). Red lines represent Re (ε), blue lines Im (ε) (ε given by (3.28)), and the green circles the eigenvalues. (a) Dynamic bifurcation when $D = D_c$, $b = 10$ and $c = -1$, $\gamma = 1.0838$. (b) Bifurcation when $\gamma = \gamma_c$, $b = 4.9341$, $c = -1.1417$. Parameters: $W_0 = -0.02$, $R = 15$, $\tau = 10$, $g = 1$, $I = 3$. For the firing rate function: $\beta = 0.8$, $\mu = 0.5$.

Let $(V(\mathbf{x},t), I(\mathbf{x},t), I_{syn}(\mathbf{x},t)) = (\overline{V}, \overline{I}, \overline{I}_{syn})$ be the homogeneous steady state, then substituting into (3.31) gives

$$\overline{V} = \frac{\overline{I}_{syn}}{\overline{R}^{-1} + r^{-1}}, \qquad \overline{I} = \frac{\overline{V}}{r}, \tag{3.33}$$

where $\overline{I}_{syn}$ is defined by

$$\begin{aligned} \overline{I}_{syn} &= \eta * W \otimes f(\overline{V}) = \int_{-\infty}^{t} \mathrm{d}t'\ \eta(t-t') \int_{\mathbb{R}^2} \mathrm{d}\mathbf{x}'\ W(\mathbf{x}-\mathbf{x}') f(\overline{V}) \\ &= \overline{W} f\left(\frac{\overline{I}_{syn}}{\overline{R}^{-1} + r^{-1}}\right). \end{aligned} \tag{3.34}$$

Now, linearisation about the steady state yields

$$C\frac{\mathrm{d}V}{\mathrm{d}t} = -\frac{V}{\overline{R}} + I_{syn} - I, \tag{3.35a}$$

$$L\frac{\mathrm{d}I}{\mathrm{d}t} = V - rI, \tag{3.35b}$$

$$I_{syn} = \gamma \eta * W \otimes V, \tag{3.35c}$$

where $\gamma = f'(\overline{V})$. Seeking solutions of the form $\mathrm{e}^{\lambda t}\mathrm{e}^{i\mathbf{k}\cdot\mathbf{r}}\left(V_0, I_0, I_{syn_0}\right)$ gives the dispersion relation

$$\varepsilon(\lambda, k) = \gamma \widetilde{W}(k)\widehat{\eta}(\lambda) - \left(C\lambda + \frac{1}{\overline{R}} + \frac{1}{L\lambda + r}\right) = 0. \tag{3.36}$$

Here $\widetilde{W}$ is given by (3.12), and $\widehat{\ }$ denotes the Laplace transform defined by

$$\widehat{\psi}(s) = \int_0^\infty \mathrm{e}^{-st}\psi(t)\mathrm{d}t, \tag{3.37}$$

such that

$$\begin{aligned}\int_{-\infty}^{t} \eta(t-t')\mathrm{e}^{\lambda t'}\mathrm{d}t' &= \mathrm{e}^{\lambda t}\int_0^\infty \left(-\alpha^2\frac{\mathrm{d}}{\mathrm{d}\alpha}\mathrm{e}^{-\alpha t'}\right)\mathrm{e}^{-\lambda t'}\mathrm{d}t' = \mathrm{e}^{\lambda t}\left(\lambda/\alpha+1\right)^{-2} \\ &= \mathrm{e}^{\lambda t}\widehat{\eta}(\lambda).\end{aligned} \tag{3.38}$$

As in §3.3.2 it is convenient to rewrite (3.36) in the form (3.26) with

$$\mathcal{G}(\nu,\omega,k) = \gamma\alpha^2\widetilde{W}(k) - \left(\left(\alpha+\nu\right)^2-\omega^2\right)\Psi_1 + 2\omega^2\left(\alpha+\nu\right)\Psi_2, \tag{3.39a}$$

$$\mathcal{H}(\nu,\omega,k) = -\left(\left(\alpha+\nu\right)^2-\omega^2\right)\omega\Psi_2 - 2\omega\left(\alpha+\nu\right)\Psi_1, \tag{3.39b}$$

and

$$\Psi_1 = C\nu + \frac{1}{\overline{R}} + \frac{L\nu+r}{\left(L\nu+r\right)^2+\left(L\omega\right)^2}, \tag{3.40}$$

$$\Psi_2 = C - \frac{L}{\left(L\nu+r\right)^2+\left(L\omega\right)^2}. \tag{3.41}$$

The chosen bifurcation parameter is C because it is unrelated to the I current dynamics and does not influence the steady state. Figure 3.16a shows the spectral picture at the bifurcation point of (3.36). As in §3.3.2, with these parameters the maximum of the dispersion curve is at $k_c = 0$ (Fig. 3.16b) and the mode that is first excited is a spatially uniform steady state. But when α or L are increased a Turing-Hopf bifurcation is found with $k_c = \max_k \widetilde{W}(k)$ (Figs. 3.16c, 3.16d). Unlike the previously studied models (§3.2, 3.3, 3.5) if the bifurcation parameter is increased (i.e., $C > C_c$) then the system is stable, whereas if it is decreased (i.e., $C < C_c$) the system is unstable.

When varying other model parameters (α and L) the critical wave-number for the Turing-Hopf bifurcation does not change. Here, the wave-number is again given by $k_c = \max_k \widetilde{W}(k)$, as in the simple model of grid cell activity (§3.2). Thus, the resonance fails to provide a mechanism for different spatial scales.

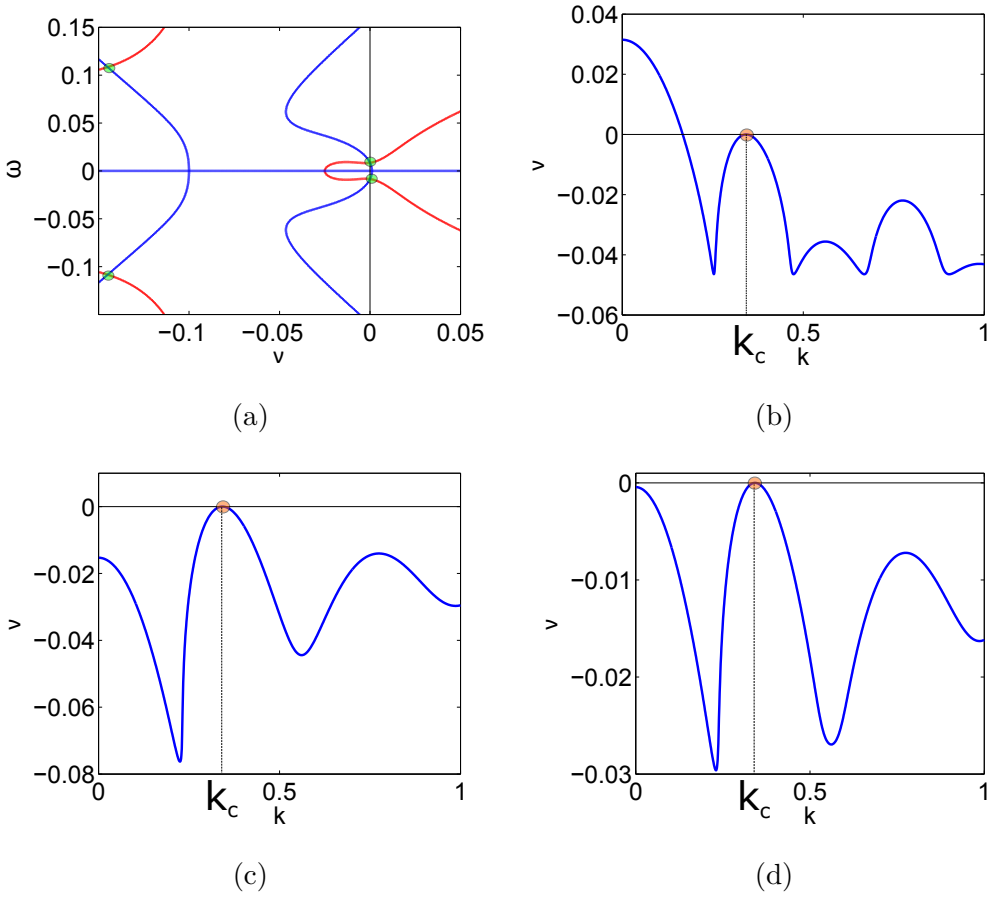


Figure 3.16. (a) Spectral picture (3.36) for the resonant model (3.31) at $C = C_c$. Red lines represent (3.39a), blue lines (3.39b) and green circles the eigenvalues (i.e, when $\nu = 0$). At the bifurcation point $C = C_c = 15.4133$ and $k = k_c = 0.3424$. Parameters: $\overline{R} = 1$, $L = 20$, $\alpha = 0.1$. Here, the connectivity is given by $W_0 = -0.02$, $R = 15$. For the firing rate funciton: $\beta = 0.8$, $\mu = 0.5$, $b = 10$, $c = 1$. (b),(c) and (d) Plots of $\mathrm{Re}\,(\lambda) = \nu$ as a function of k for different set of parameters. Orange circle represents the value of k_c given by the `fsolve` function in MATLAB. In (b) the same set of parameters as in (a) are used, and it shows that the computed value of k_c does not correspond to the critical value. An accurate solution for the bifurcation point is obtained when varying the parameters α and L. The new parameters are: in (c) $\alpha = 0.25$, $C_c = 51.0409$, $L = 20$; and in (d) $\alpha = 0.1$, $C_c = 90.3996$, $L = 40$.

3.6 MODEL WITH DENDRITES

So far neurons have been treated as point objects, disregarding some of their various components. In particular, dendrites are known to play a critical role in integrating synaptic inputs [15]. Here, the previous models are expanded to include dendritic processing (as described in §2.3), and axo-dendritic connections as in [8]. The dendritic tree can act as a spatio-temporal filter: it receives temporal inputs from different neurons at fixed points, and these inputs generate a response at the soma which depends on the time for signals to diffuse along the branches of the tree [47]. Work by Mainen and Sejnowski supports the hypothesis that the difference in firing patterns between neocortical neurons may be caused by their dendritic morphology [48].

The proposed model is an effective single population model in two (somatic) spatial dimensions, and a further dimension representing position along a (semi-infinite) dendritic cable. The firing rate in the somatic (cell body) layer is taken to be a smooth function of the cable voltage at the soma, which is in turn determined by the spatio-temporal pattern of synaptic currents on the cable. The voltage $V(\xi, \mathbf{x}, t)$ at position $\xi \geq 0$ along the semi-infinite passive cable with somatic coordinate $\mathbf{x} \in \mathbb{R}^2$ can then be written as in §2.3:

$$\frac{\partial V}{\partial t} = -\frac{V}{\tau} + D\frac{\partial^2 V}{\partial \xi^2} + \rho I_{syn}(\xi, \mathbf{x}, t). \tag{3.42}$$

Here, $I_{syn}(\xi, \mathbf{x}, t)$ is given by (3.32) and $\rho > 0$ is a constant of proportionality representing the strength of the synaptic input. Unlike previous sections, $W(\xi, \mathbf{x})$ describes the axo-dendritic connectivity and it is assumed that it can be decomposed in the product form

$$W(\xi, \mathbf{x}) = \delta(\xi - \xi_0)W(\mathbf{x}), \tag{3.43}$$

where ξ_0 is a fixed distance from the soma. Of importance in what follows will be the "drive" at the soma, $V(\xi = 0, \mathbf{x}, t)$, which is denoted henceforth by $h(\mathbf{x}, t)$. For no flux boundary conditions, $\partial V(\xi, \mathbf{x}, t)/\partial \xi|_{\xi=0} = 0$, and assuming vanishing initial data, the solution to (3.42) at $\xi = 0$ becomes

$$h(\mathbf{x}, t) \equiv V(\xi = 0, \mathbf{x}, t) = \rho(G \otimes I_{syn})(\xi = 0, \mathbf{x}, t); \qquad G = 2G_\infty. \tag{3.44}$$

Here the operator $\otimes$ denotes spatio-temporal convolution over the (ξ, t) coordinates and $G_\infty(\xi, t)$ is given by (2.40). Note that in obtaining (3.44), the result of the Green's function (between two points ξ and ξ') for the semi-infinite cable with no flux boundary conditions can be written as $G_\infty(\xi - \xi', t) + G_\infty(\xi + \xi', t)$, has been used. Therefore,

$$h(\mathbf{x}, t) = \rho \int_{-\infty}^{t} \mathrm{d}s' \int_0^\infty \mathrm{d}\xi'\ G(\xi', t - s') I_{syn}(\xi', \mathbf{x}, s'). \tag{3.45}$$

3.6.1 *Linear Stability Analysis*

The spatially homogeneous steady state $h(\mathbf{x}, t) = \overline{h}$ of (3.45) is given by

$$\begin{aligned} \overline{h} &= 2\rho \overline{W} f(\overline{h}) \int_0^\infty \mathrm{d}s \frac{1}{\sqrt{4\pi D s}} \mathrm{e}^{-s/\tau} \mathrm{e}^{-\xi_0^2/(4Ds)} \\ &= \rho \overline{W} f(\overline{h}) \left(\frac{\sqrt{\tau} \mathrm{e}^{-\sqrt{\frac{\xi_0^2}{D\tau}}}}{\sqrt{D}} \right). \end{aligned} \tag{3.46}$$

Note that this steady state depends on all the parameters in the model, which makes the analysis difficult. One way to avoid this difficulty is to add an external drive I_{ext} to the synaptic input, thus $I_{syn} \to I_{syn} + I_{ext}$ in (3.42) where

$$I_{ext}(\xi) = -f(\overline{h}) \int_{\mathbb{R}^2} \mathrm{d}\mathbf{x}'\ W(\xi, \mathbf{x} - \mathbf{x}'). \tag{3.47}$$

Substitution in (3.45) gives

$$h(\mathbf{x}, t) = \rho \int_{-\infty}^{t} \mathrm{d}s' \int_0^\infty \mathrm{d}\xi'\ G(\xi', t - s') \left(I_{syn}(\xi', \mathbf{x}, s') + I_{ext}(\xi') \right), \tag{3.48}$$

that admits only the trivial spatially homogeneous steady state, $\bar{h} = 0$. Because I_{ext} is constant with respect to the somatic coordinate $\mathbf{x}$ and time, the addition of this current to the synaptic input does not affect the analysis since it is dropped in the linearisation about $\overline{h}$, which yields

$$h(\mathbf{x}, t) = \rho\gamma \int_{-\infty}^{t} \mathrm{d}s' \int_0^\infty \mathrm{d}\xi' G(\xi', t - s') \int_{-\infty}^{s'} \mathrm{d}s \eta(s' - s) \int_{\mathbb{R}^2} \mathrm{d}\mathbf{x}' W(\xi', \mathbf{x} - \mathbf{x}') h(\mathbf{x}', s), \tag{3.49}$$

where $\gamma = f'(\overline{h})$. Seeking solutions of the form $h(\mathbf{x}, t) = h_0 \mathrm{e}^{\lambda t} \mathrm{e}^{i\mathbf{k}\cdot\mathbf{x}}$ and substituting in (3.49), the following spectral equation is obtained:

$$1 = \rho\gamma \widehat{\eta}(\lambda) \widetilde{W}(k) \widehat{G}(\xi_0, \lambda), \tag{3.50}$$

where

$$\widehat{G}(\xi,\lambda)=\int_0^\infty G(\xi,s)\mathrm{e}^{-\lambda s'}\mathrm{d}s'=\frac{\mathrm{e}^{-\varphi(\lambda)\xi}}{D\varphi(\lambda)},\qquad \varphi^2(\lambda)=\frac{1/\tau+\lambda}{D}, \tag{3.51}$$

is the Laplace transform of G (defined by (3.44); see Appendix D for details). The dispersion relation is given by

$$\varepsilon(\lambda,k)=(\lambda+\alpha)^2-\alpha^2\rho\gamma\widetilde{W}(k)\widehat{G}(\xi_0,\lambda)=0. \tag{3.52}$$

3.6.2 *Numerical Scheme for the Dispersion Relation*

Due to its complexity, it is problematic to solve the dispersion relation (3.52) using MATLAB's `fsolve` function in the system of equations described in §3.3.2. Thus, a simple numerical scheme is used by taking $\lambda=\nu+i\omega$ and $\varepsilon(\nu+i\omega,k)=\mathcal{G}(\nu,\omega,k)+i\mathcal{H}(\nu,\omega,k)$, in which

1. A value of the bifurcation parameter is chosen and MATLAB's `fsolve` function is used to find ν and ω satisfying $\mathcal{G}(\nu,\omega,k)=0=\mathcal{H}(\nu,\omega,k)$ for a range of k. Here, the selected bifurcation parameter is ρ due its linearity in (3.52).

2. A maximum value of $\nu(k)$ is found and if it is less than 0, then the bifurcation parameter is less than its critical value and therefore it is increased.

3. The process in 2) is repeated until $\max_k(\nu)\geq 0$, and the first value of ρ for which this occurs is the critical value of the bifurcation parameter.

Although the solution accuracy to the dispersion relation under this numerical scheme depends on the choice of increments in ρ and the selected mesh for the range of k, its construction gives a good approximation for the critical wave-number.

Figure 3.17 shows the spectral pictures at the bifurcation point with different values of ξ_0. When $\xi_0=2$ there is a static Turing bifurcation (Fig. 3.17a) whereas when $\xi_0=2.5$, $k_c=0$ and $\omega_c\neq 0$ resulting in a homogeneous dynamic

bifurcation, hence the system will evolve to a coherent network oscillation with temporal frequency ω_c, as shown in Figure 3.17b. Moreover, the exact value when the bifurcation changes can be determined by performing the continuation of ξ_0 (Fig. 3.18) or other model parameters: D, τ and α (data not shown).

These numerical continuation results show that when there is a static patterned bifurcation, the critical wave-number is given by $\max_k \widetilde{W}(k)$ as in previous models. Hence, dendrites and a separable axo-dendritic connectivity fail to provide a mechanism to change the size and spacing of the patterns.

3.6.3 *Model with Correlated Axo-Dendritic Connections*

Here, the effects of variation in the spatial locations of synapses along the dendritic tree are investigated by using a correlated weight distribution as in Bressloff and Coombes [47]. This weight distribution is based on the idea that synapses tend to be located further away from the soma in the dendritic tree as the separation between neurons increases, and that they excite the postsynaptic neuron at the contact point. Thus, the axo-dendritic connectivity function is given by

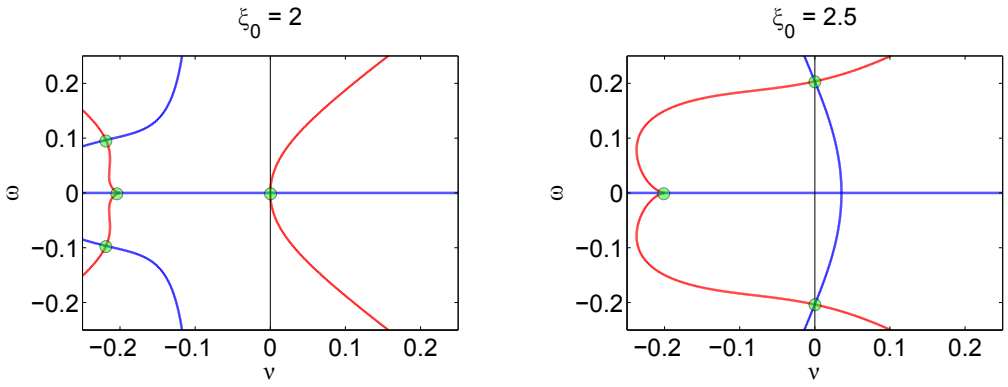
$$W(\xi, \mathbf{x}) = \delta(\xi - \xi_0 - \kappa\,|\mathbf{x}|)W(\mathbf{x}). \tag{3.53}$$

Seeking solutions of the form $h(\mathbf{x}, \xi) = h_0 \mathrm{e}^{\lambda t}\mathrm{e}^{i\mathbf{k}\cdot\mathbf{x}}$ in the linearised system (3.49) gives

$$\begin{aligned}
h_0\mathrm{e}^{\lambda t}\mathrm{e}^{i\mathbf{k}\cdot\mathbf{x}} &= \vartheta(\lambda)\int_{-\infty}^{t}\mathrm{d}s'\int_0^{\infty}\mathrm{d}\xi' G(\xi', t-s')\mathrm{e}^{\lambda s'}\int_{\mathbb{R}^2}\mathrm{d}\mathbf{x}' W(\xi', \mathbf{x}-\mathbf{x}')\mathrm{e}^{i\mathbf{k}\cdot\mathbf{x}'} \\
&= \vartheta(\lambda)\frac{\mathrm{e}^{\lambda t}\mathrm{e}^{i\mathbf{k}\cdot\mathbf{x}}}{D\varphi(\lambda)}\int_0^{\infty}\mathrm{d}\xi'\mathrm{e}^{-\varphi(\lambda)\xi'}\delta(\xi'-\xi_0-\kappa\left|\mathbf{x}'\right|)\int_{\mathbb{R}^2}\mathrm{d}\mathbf{x}' W(\mathbf{x}')\mathrm{e}^{i\mathbf{k}\cdot\mathbf{x}'} \\
&= \vartheta(\lambda)\frac{\mathrm{e}^{\lambda t}\mathrm{e}^{i\mathbf{k}\cdot\mathbf{x}}\mathrm{e}^{-\varphi(\lambda)\xi_0}}{D\varphi(\lambda)}\int_{\mathbb{R}^2}\mathrm{d}\mathbf{x}' W_0 H(R-\left|\mathbf{x}'\right|)\mathrm{e}^{i\mathbf{k}\cdot\mathbf{x}'}\mathrm{e}^{-\varphi(\lambda)\kappa|\mathbf{x}'|} \\
&= \vartheta(\lambda)\frac{\mathrm{e}^{\lambda t}\mathrm{e}^{i\mathbf{k}\cdot\mathbf{x}}\mathrm{e}^{-\varphi(\lambda)\xi_0}}{D\varphi(\lambda)}2\pi W_0\int_0^{R}J_0(rk)r\mathrm{e}^{-\varphi(\lambda)\kappa r}\mathrm{d}r,
\end{aligned} \tag{3.54}$$

where $\vartheta(\lambda) = \rho\gamma h_0\widehat{\eta}(\lambda)$. Therefore

$$1 = \rho\gamma\widehat{\eta}(\lambda)\frac{\mathrm{e}^{-\varphi(\lambda)\xi_0}}{D\varphi(\lambda)}\Omega(\lambda, k), \tag{3.55}$$



(a) (b)

Figure 3.17. Spectral picture (3.52) for the model with dendrites (3.42) and different values of ξ_0. Here, red line represents the real part of (3.52) and the imaginary part is represented with a blue line. Green circles are the eigenvalues. Parameters: $D = 1$, $\tau = 5$, $\alpha = 0.1$. Hereinafter, unless stated otherwise, the connectivity is given by $W_0 = -0.02$, $R = 15$ and for the firing rate funciton: $\beta = 0.8$, $\mu = 0.5$, $b = 10$, $c = -1$. (a) Static bifurcation with $k_c = 0.3424$, $\omega_c = 0$, $\rho_c = 435.94$. (b) Dynamic bifurcation with $k_c = 0$, $\omega_c = 0.2035$, $\rho_c = 496.11$.

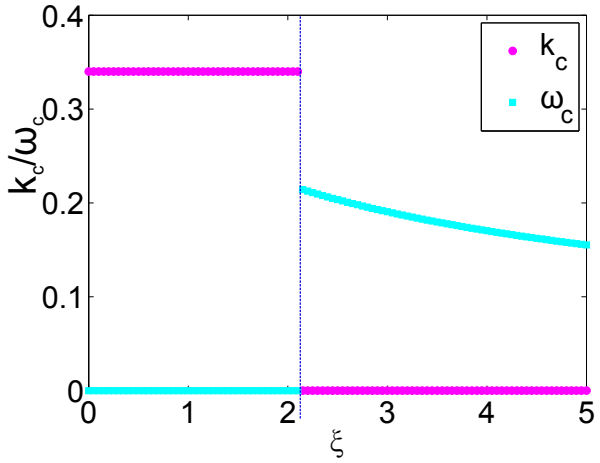


Figure 3.18. Continuation of ξ_0 at the bifurcation point for the model with dendrites (3.42). Other parameters are as in Figure 3.17. The blue dotted line represents the value of ξ_0 for which the bifurcation changes from a static patterned ($\omega_c = 0,\ k_c \neq 0$) to a dynamic homogeneous one ($\omega_c \neq 0,\ k_c = 0$).

with

$$\Omega(\lambda, k) = 2\pi W_0 \int_0^R J_0(rk) r \mathrm{e}^{-\varphi(\lambda)\kappa r} \mathrm{d}r. \tag{3.56}$$

The dispersion relation is given by

$$\varepsilon(\lambda, k) = (\lambda + \alpha)^2 - \alpha^2 \rho\gamma \frac{\mathrm{e}^{-\varphi(\lambda)\xi_0}}{D\varphi(\lambda)} \Omega(\lambda, k) = 0. \tag{3.57}$$

In order to find the critical wave-number, the numerical scheme given in §3.6.2 is implemented with ρ as the bifurcation parameter. Figure 3.19 shows the continuation with other parameters for the critical wave-number (magenta) and critical temporal frequency (cyan). Note that the model can either exhibit a static patterned bifurcation or a dynamic homogeneous one depending on the value of other parameters (κ in Fig. 3.19a, τ in Fig. 3.19b and D in Fig. 3.19c), as in the model without the axo-dendritic connections. As before, there is no effect in the critical wave-number when varying α or ξ_0. However, there is a small change in the wave-number when varying the parameters D, τ or κ (see Fig. 3.20). Importantly, the variation in the critical wave-number is due to their nonlinear relation with k_c in the dispersion relation via $\Omega(\lambda, k)$ (see (3.57)).

Adding axo-dendritic connections changes the relationship between k_c and the dispersion relation, resulting in a different critical wave-number than in the previous model (given by $\max_k \widetilde{W}_k$). Although the axo-dendritic connections gave a mechanism for changing the pattern scale, such variation is insignificant when compared with the variation obtained when changing the radius of connectivity in the simple model of grid cell activity (see Fig. 3.11). Note that the variation in scale when varying different parameters has the same magnitude (Fig. 3.19). Thus, when adding the head-direction and velocity inputs (from the animal trajectory) to the model, the change in size of the resulting firing fields would not resemble the change seen in experimental data.

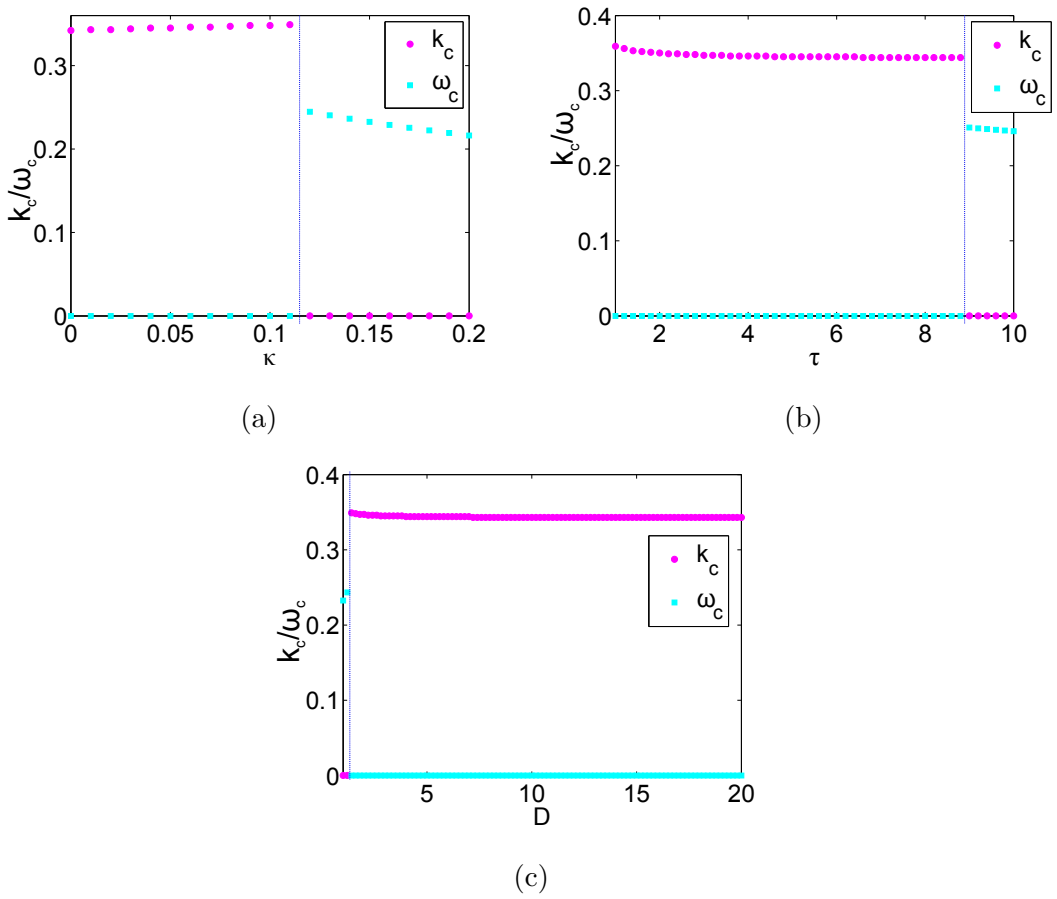
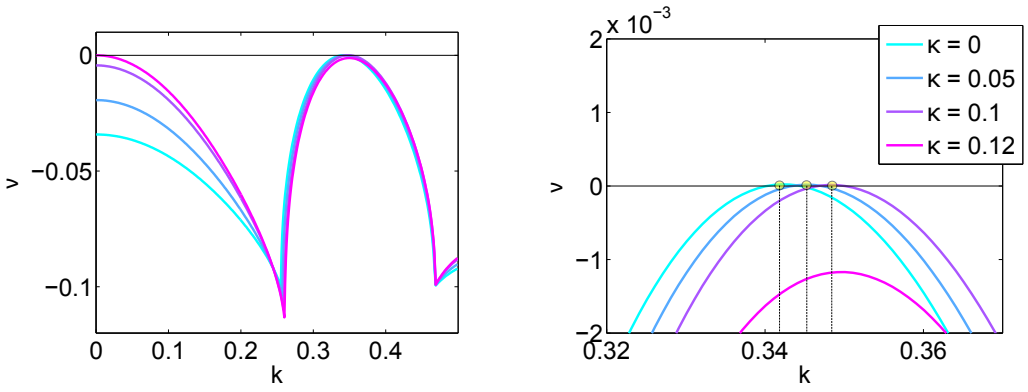


Figure 3.19. Parameter continuation for the model with correlated Axo-Dendritic connections (3.53). Dotted blue line indicates the parameter value for which the type of bifucation switches. (a) shows the continuation on κ, (b) on τ and (c) on D. Parameters: $\alpha = 0.1$, $\tau = 5$, $D = 1$, $\xi_0 = 0.2$, for (b) $\kappa = 0.1$ and for (c) $\kappa = 0.15$.



(a) (b)

Figure 3.20. Plot of Re (λ) as a function of k in the dispersion relation (3.57) for the model with correlated axo-dendritic connections. Different colours denote different vales of κ. (b) Zoom of (a), yellow circles represent the critical values of k for different κ. Parameters: $\alpha = 0.1$, $\tau = 5$, $D = 1$, $\xi_0 = 0.2$.

3.6.4 *Gaussian Axo-Dendritic Connections*

In §3.6.3 there was a small change in the critical wave-number when varying model parameters. Here, with the aim of having a more realistic description of the synapses, it is assumed that the synapse contact point on the postsynaptic neuron has a delimited area and that the strength of the input varies around the centre of this area. Therefore a smooth version of the dendritic connection is introduced and it is investigated if such a connection increases the range of k_c that is obtained when performing continuation of the parameters. The smooth version of the dendritic connection is defined as

$$W(\xi, \mathbf{x}) = \Lambda(\xi - \xi_0 - \kappa\,|\mathbf{x}|)W(\mathbf{x}), \qquad \Lambda(\xi) = \frac{\mathrm{e}^{-\xi^2/\sigma^2}}{\sqrt{2\pi\sigma^2}}. \tag{3.58}$$

Here σ can be interpreted as the width of the connection point. Thus for $\sigma \to 0$ the connectivity with the δ-function given by (3.53) is recovered.

The dispersion relation for this type of connectivity is given by

$$\varepsilon(\lambda, k) = (\lambda + \alpha)^2 - \alpha^2\rho\gamma\frac{\mathrm{e}^{-\varphi(\lambda)\xi_0}\mathrm{e}^{\varphi^2(\lambda)\sigma^2/4}}{2\sqrt{2}D\varphi(\lambda)}\Omega(\lambda, k) = 0, \tag{3.59}$$

where

$$\Omega(\lambda,k)=2\pi W_0\int_0^R \operatorname{erfc}\left(\frac{\varphi(\lambda)\sigma^2-2\left(\xi_0+\kappa r\right)}{2\sigma}\right)J_0(rk)\mathrm{e}^{-\kappa r\phi(\lambda)}r\mathrm{d}r. \quad (3.60)$$

Here, $\operatorname{erfc}(z),\ \forall z\in\mathbb{C}$, is the complementary error function defined by

$$\operatorname{erfc}(z)=1-\operatorname{erf}(z)=\frac{2}{\sqrt{\pi}}\int_z^\infty \mathrm{e}^{-t^2}\mathrm{d}t, \quad (3.61)$$

where the error function, $\operatorname{erf}(z)$, is defined by

$$\operatorname{erf}(z)=\frac{2}{\sqrt{\pi}}\int_0^z \mathrm{e}^{-t^2}\mathrm{d}t, \quad (3.62)$$

For the numerical scheme in §3.6.2, the freely available code in [49] is used since MATLAB's error function is only defined for real numbers.

There is a small variation in k_c when taking ρ as the bifurcation parameter and varying σ (Fig. 3.21). However, this variation in the wave-number has the same magnitude with that obtained in the model where the contact point is a δ-function. Moreover, when comparing with the results given in §3.6.3 and taking a small σ it can be seen that the resulting wave-number is the same (Fig. 3.22). Although the results are the same, this Gaussian version of the connection is used hereinafter because it will facilitate the simulations in §3.7.

3.6.5 *Smooth Version of Inverse Top Hat Somatic Connectivity*

In the last section a new model parameter, the width of the connection area at the dendritic tree σ, is introduced when the δ-function dendritic connection is replaced by a Guassian one. This provides a mechanism for changing the pattern scale, but the variation in the scale is small. Now, it is investigated if a smooth somatic connection kernel could give a greater variation. Thus, the somatic connection is modified to

$$W(\mathbf{x})=\frac{W_0}{2}\left(\tanh\left(\varsigma\left(R-|\mathbf{x}|\right)\right)+\tanh\left(\varsigma\left(R+|\mathbf{x}|\right)\right)\right). \quad (3.63)$$

As shown in Figure 3.23a the smoothness of the connectivity depends on ς. The dispersion relation is the same as in (3.59), but now $\Omega(\lambda,k)$ is given by

$$\Omega(\lambda,k)=2\pi\int_0^\infty \operatorname{erfc}\left(\frac{\varphi(\lambda)\sigma^2-2\left(\xi_0+\kappa r\right)}{2\sigma}\right)J_0(rk)\mathrm{e}^{-\kappa r\phi(\lambda)}W(r)r\mathrm{d}r. \quad (3.64)$$

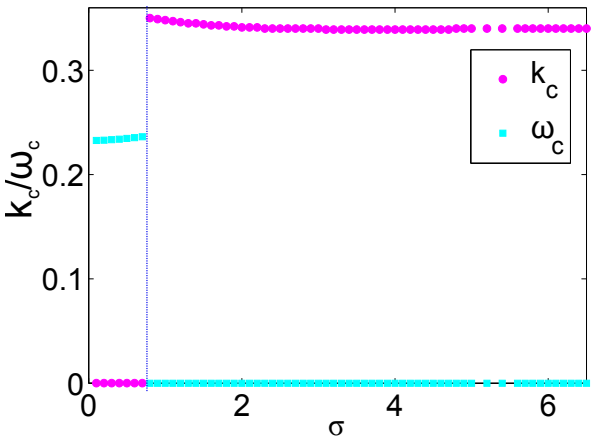
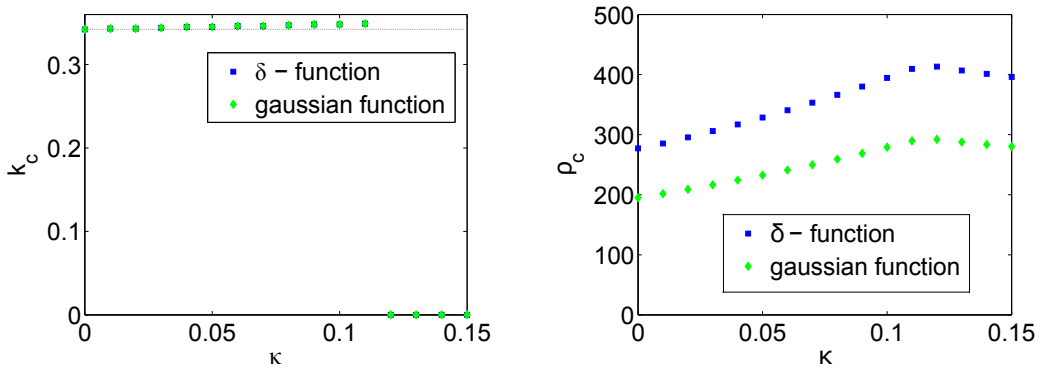


Figure 3.21. Continuation in σ for the model with Gaussian axo-dendritic connections (3.58). Blue dotted line signals the value of σ where the bifurcation changes from a dynamical homogeneous to a static patterned. Parameters: $D = 1$, $\tau = 5$, $\alpha = 0.1$, $\kappa = 0.15$, $\xi_0 = 0.2$



(a) (b)

Figure 3.22. Comparison between the different types of dendritic connections (Guassian (3.58) vs. δ-function (3.53)) in the model with correlated axo-dendritic connections. (a) shows the critical wave-number whereas (b) the critical bifurcation parameter. Here $\alpha = 0.1$, $\tau = 5$, $D = 1$, $\xi_0 = 0.2$, $\sigma = 0.1$.

As shown in Figure 3.23b, for $\varsigma = 0.15$ (cyan dots) the range of critical wave-numbers resulting from the continuation of a parameter, in this case κ, is slightly bigger than when using a inverse Top Hat connectivity described in (3.5) (black dots), but as ς is further decreased the range of wave-numbers decreases until for $\varsigma \leq 0.1$, $k_c = 0 \; \forall k$.

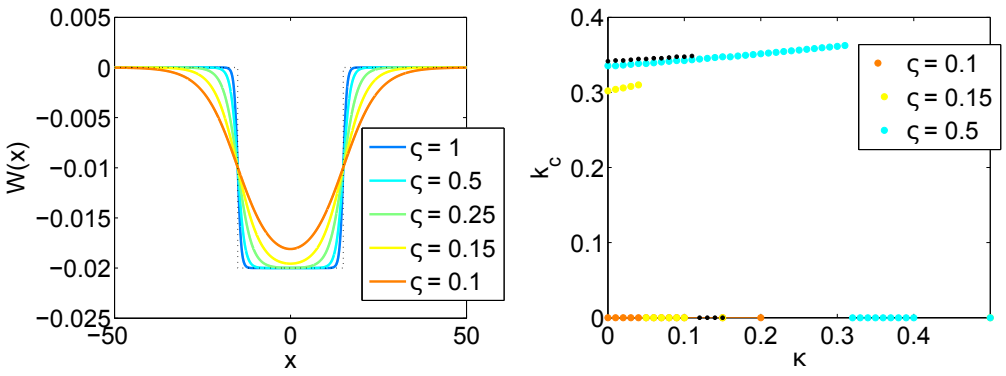
The range of pattern wavelengths found when taking a smooth representation of the somatic connectivity is not sufficient to illustrate the hypothesised different scales in grid cell activity along the dorsoventral axis of the MEC by computational models of grid cells [26, 39]. Moreover, the patterns disappear when increasing the smoothness of the connectivity.

3.6.6 *Inverse Mexican Hat Somatic Connectivity*

It was shown in §3.6.5 that a wider range of critical wave-numbers can be obtained when using a smooth somatic connectivity, though such a range is insuffient to achieve a large difference in the scale of emerging patterns. Now, the idea that grid cells may also receive a positive input from neurons (I in (3.4)) is incorporated by proposing an inverse Mexican Hat connectivity such as that employed in [39]. Such a connectivity exhibits a short range inhibition and a long range excitation (Fig. 3.25a). In this situation, inhibitory cells connect directly to the dendritic tree of excitatory cells, whereas excitatory cells connect via the axon to both excitatory and inhibitory cells as shown in Figure 3.24. Here the somatic connectivity is given by

$$W(\mathbf{x}) = \alpha_1 \mathrm{e}^{-|\mathbf{x}|\sigma_1} + \alpha_2 \mathrm{e}^{-|\mathbf{x}|\sigma_2}, \tag{3.65}$$

and the dispersion relation is the same as in (3.59), with Ω as in (3.64). Figure 3.25b shows that with this type of connectivity the critical wave-number varies with κ. Furthermore, this range of k_c increases when the difference between the inhibition and excitation of the proposed connectivity increases. Unlike the previous models with correlated axo-dendritic connections, when κ is small there is a homogeneous static bifurcation (i.e $k_c = \omega_c = 0$) and a Turing-Hopf



(a) (b)

Figure 3.23. (a) Smoothed Top Hat function as in (3.63) for different values of ς. (b) Continuation of κ in the model with Gaussian axo-dendritic connections with the somatic connectivity in (a). Different colours represent different values of ς and black dots correspond to the simulation with a inverse Top Hat connectivity as in Figure 3.22a.

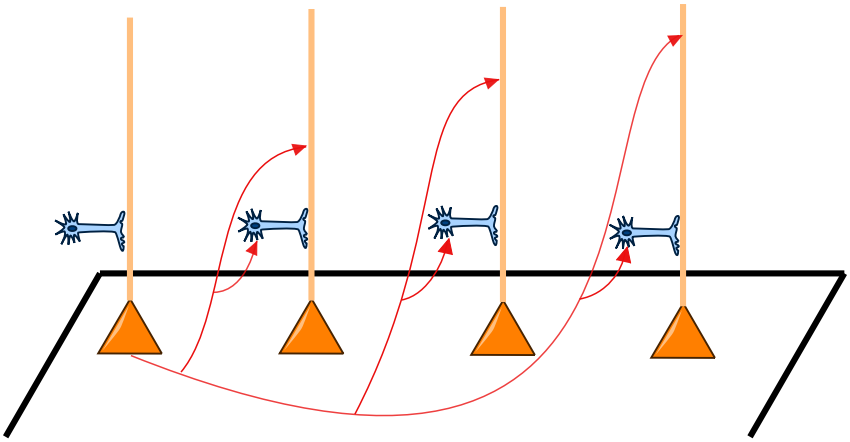


Figure 3.24. Diagram of a neural tissue consisting of two types of neurons: pyramidal (orange) and stellate (blue). Stellate cells inhibit pyramidal cells that excite each other in their dendritic tree (orange line) depending on the separation in the soma (triangle). Note that the pyramidal neurons axon (red line) branch to contact different neurons whereas stellate cell just inhibit one pyramidal neuron.

bifurcation when κ increases. If κ is increased further, then the bifurcation is homogeneous dynamic (i.e $k_c = 0$, $\omega_c \neq 0$). Hence, with this type of connectivity it is possible to have temporal oscillations with frequency $\omega_c \neq 0$.

Figure 3.25c shows that the range of critical wave-numbers is similar when varying the time constant, τ, and diffusion constant, D. In these cases, the bifurcations are always dynamic and the critical wave-number converges to a fixed value when the parameters increase. Thus, this model gives a mechanism for different pattern wavelengths. Unfortunately, there is no experimental evidence that the time or diffusion constant changes along the dorso-ventral axis of the MEC. Likewise, it is improbable that κ changes along this axis.

3.7 RESONANT MODEL WITH AXO-DENDRITIC CONNECTIONS

In §3.6.6 a mechanism for different pattern wavelengths was found, but lacks experimental support. Therefore, a membrane current I is added, as in §3.5, based on the fact that temporal frequency of subthreshold oscillations changes across the dorso-ventral axis of MEC [46]. Moreover, recent experimental data from layer II stellate cells of MEC shows that the time constant of I_h currents changes between dorsal and ventral neurons [50]. For numerical convenience, the model is now one dimensional in the soma and dendritic coordinate, with $V = V(\xi, x, t)$, $x, \xi \in \mathbb{R}$. This model can be described by

$$\frac{\partial V}{\partial t} = -\frac{V}{\tau} + D\frac{\partial^2 V}{\partial \xi^2} - \frac{1}{C}\left(I - gI_s\right), \tag{3.66a}$$

$$L\frac{\mathrm{d}I}{\mathrm{d}t} = -rI + V. \tag{3.66b}$$

As in §3.6, $I_{syn} \longrightarrow I_{syn} + I_{ext}$ in order to facilitate linear analysis. Travelling waves (Fig. 3.26) were observed when simulating the system. Therefore, a linear stability analysis is performed in order to validate the observed wave-number.

Taking $h = V(0, x, t)$ as in §3.6.1, linearising via small perturbations of the form $h(\xi, x, t) = h_0 \mathrm{e}^{\lambda t}\mathrm{e}^{ikx}$ gives the dispersion relation

$$\varepsilon(\lambda, k) = 1 - \frac{g\gamma\widehat{\eta}(\lambda)\mathrm{e}^{\left(\frac{\sigma^2\varphi^2(\lambda)}{4} - \varphi(\lambda)\xi_0\right)}}{\sqrt{2}CD\varphi(\lambda)}\mathrm{Re}\left(\Omega_1(k) + \Omega_2(k)\right) = 0, \tag{3.67}$$

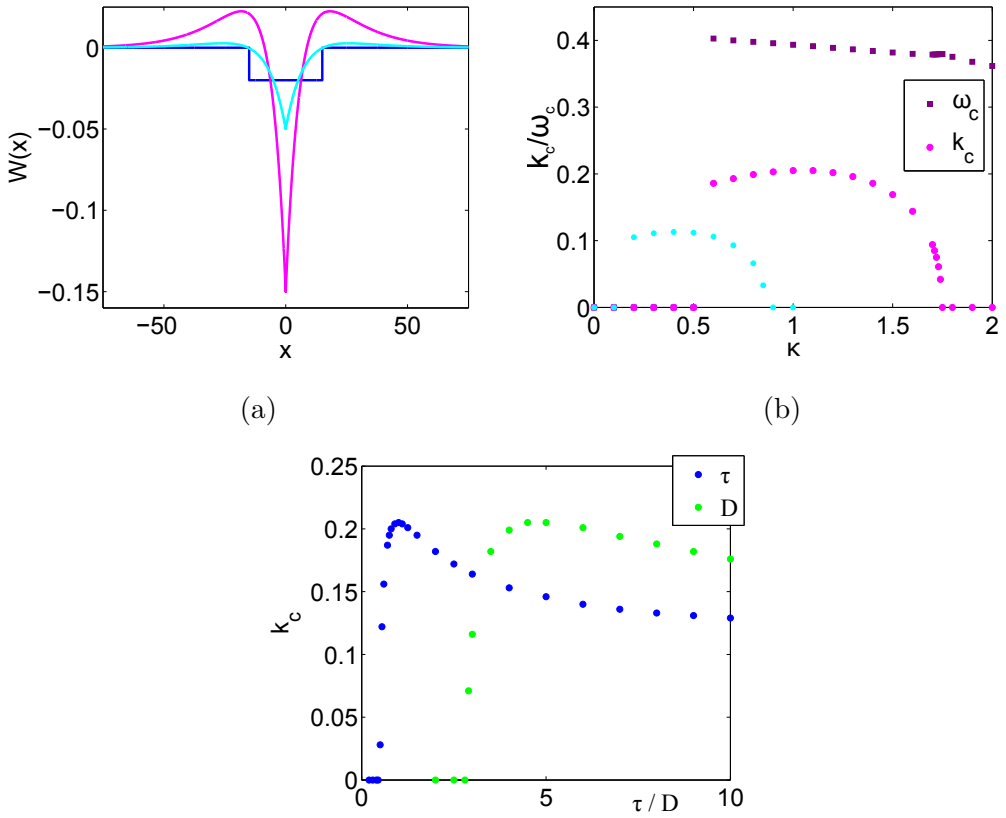


Figure 3.25. (a) Inverse Mexican Hat somatic connectivity (3.65) for various parameter choices. Cyan: $\alpha_1 = 0.05$, $\alpha_2 = -0.1$, $\sigma_1 = 0.075$, $\sigma_2 = 0.12$. Magenta: $\alpha_1 = 0.15$, $\alpha_2 = -0.3$, $\sigma_1 = 0.07$, $\sigma_2 = 0.15$. Blue line indicates the inverse Top Hat connectivity. (b) Continuation in κ of a model with Gaussian axo-dendritic connections and somatic connectivities as in (a). Circles represents the resulting critic wave-number for different connectivity parameters (colour-coded as in (a)). Purple squares represent the resulting critical temporal frequency in the system with parameters as the ones in magenta. (c) Continuation in τ and D. Here, $\kappa = 0.15$ and the somatic connectivity parameters are the same as the ones represented in magenta in (a). Other parameters as in Figure 3.22.

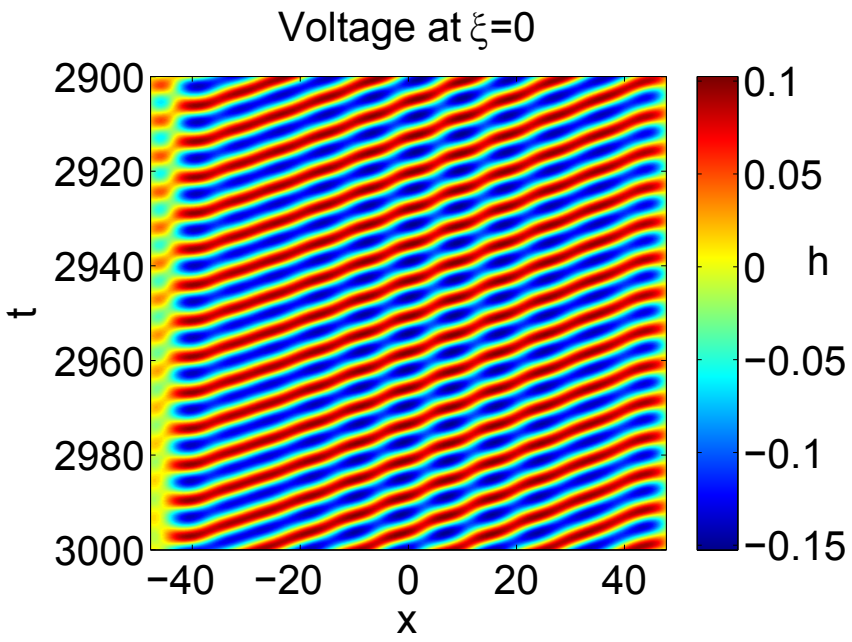


Figure 3.26. Activity in the soma $V(0,x,t) = h(x,t)$ resulting from the simulation of the resonant model with axo-dendritic connections (3.66) (See Appendix E for details). Parameters: $\alpha_1 = 1$, $\alpha_2 = -4$, $\sigma_1 = 0.5$, $\sigma_2 = 1$, $\tau = 1$, $D = 5$, $\alpha = 0.1$, $\sigma = 1$, $\kappa = 0.5$, $L = C = r = 1$, $g = 163620$, $\xi_0 = 2$.

where $\widehat{\eta}(\lambda)$ is defined in (3.38) and $\gamma = f'(0)$. Following the calculations of Appendix D gives

$$\varphi^2(\lambda) = \frac{1}{D}\left(\lambda + \frac{1}{\tau} + \frac{1}{C\left(L\lambda + r\right)}\right). \tag{3.68}$$

Here

$$\Omega_j(k) = \frac{\alpha_j}{(\sigma_j + \varphi(\lambda)\kappa + ik)}\left(\operatorname{erfc}\left(\phi\right) + \mathrm{e}^{-\phi^2}\mathrm{e}^{\psi^2}\operatorname{erfc}\left(\psi\right)\right), \tag{3.69}$$

where

$$\phi = \frac{\varphi(\lambda)\sigma^2 - 2\xi_0}{2\sigma}, \qquad \psi = \frac{\left(\sigma_j + \frac{2\kappa\xi_0}{\sigma^2} + ik\right)\sigma}{2\kappa}. \tag{3.70}$$

Here g is the bifurcation parameter. Importantly, the analysis predicts the critical wave-number obtained in the simulations (see Appendix E for details). Figure 3.27a shows the critical wave-number when varying κ. Interestingly, ω_c remains constant throughout this variation. Thus, the bifurcation changes from patterned to homogeneous as κ increases and it is always dynamic. When varying τ and D (Fig. 3.27b) the system goes from a homogeneous state to

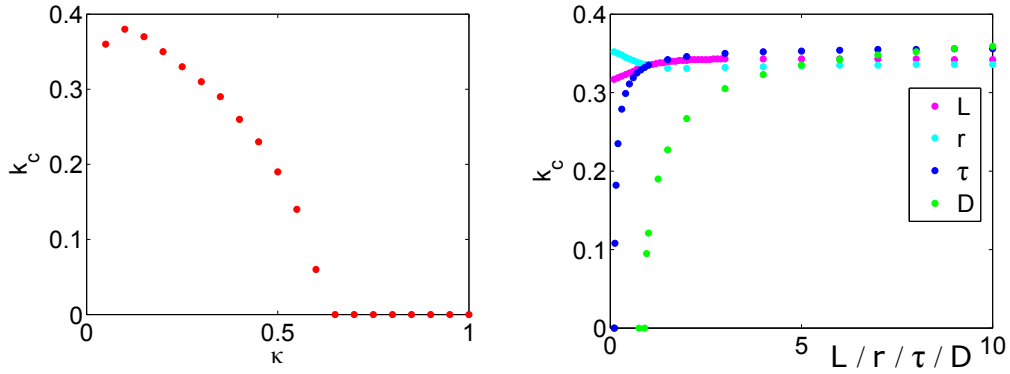
a patterned dynamic bifurcation. As in the model without resonance in §3.6, the critical wave-number converges to a fixed value when the time or diffusion constants are increased. On the other hand, when varying the parameters r and L involved in the current dynamics, the change in k_c is minimal compared to that observed under variation of κ, τ and D.

Because L is proportional to the time scale of the current (see §2.2.2 for details), a significant change in the critical wave-number was expected when varying that parameter. Nevertheless, the change obtained is insufficient to show vastly different spatial scales in the generated patterns. Therefore, the resonant model with axo-dendritic connections presented in this section was not able to relate the experimental evidence of change in I_h current time scale across the MEC with different pattern spatial scales.

3.8 CONCLUSIONS

In order to find a mechanism that relates the different scales in the grid cell's firing fields with model parameters, a simple model of grid cell activity (based on the Couey *et al.* model [26]) was proposed and simulated. The spatial scale was related to the model parameters by performing a linear stability analysis. As shown in [26], this spatial scale is associated with the difference in size of the grid cell firing fields. Unfortunately, the model in §3.2 only allows the control of the firing field spatial scale via the neural connectivity distance, R. This parameter is the one used to change the spatial scale in grid cell firing in the models proposed by Couey [26], and Burak and Fiete [39]. Such structural changes in connectivity within the MEC are not physiologically reasonable. Therefore, in this chapter, a series of physiologically-based components were added to the basic model in (3.4) with the aim of uncovering a mechanism for grid cell dynamics that accounts for the difference in spatial scale independent of the connectivity.

Firstly, in §3.3 adaptation was added to the simple model and a Turing-Hopf bifurcation was found. However, there was no change in the critical wave-



(a) (b)

Figure 3.27. Continuation for different parameters in the resonant model with axo-dendritic connections (3.66). Plots show the resulting critical wave-number when varying κ (in (a)), L, r, τ and D (colour-coded in (b)). Parameters as in Figure 3.26.

number when varying other model parameters unrelated to the connectivity. Thus, second order synapses was implemented in the model (§3.3.2), together with adaptation, but similar results were obtained. Next, delays were included in the simple model (§3.4) and a static patterned bifurcation was found; however, as in the simple model, the spatial scale only depended on connectivity parameters. Then, in §3.5 current dynamics were included into the simple model while conserving its connectivity. A Turing-Hopf bifurcation was found but the spatial scale was controlled only by the radius connectivity.

Consequently, in §3.6 dendritic processing was added to the simple model. It was found, by performing continuation of model parameters, that the bifurcation changes from a static patterned to a dynamic homogeneous one. Small changes in the wave-number were observed when assuming axo-dendritic connections and varying model parameters, such as the time and diffusion constants, and the strength of the axo-dendritic connection. After noticing that the range of wave-numbers is slightly increased when taking a smooth version of the somatic connectivity in the neuron, an inverted Mexican Hat connectivity was suggested, as in [39]. Importantly, a significant change in the critical wave-

number was found when varying κ, τ or D. However, there is no experimental evidence that supports the variation of such parameters in grid cells.

Finally, in view of experimental evidence that states that the time constant of the I_h current varies along the dorso-ventral axis of the MEC [50], axo-dedritic connections with inverse Mexican Hat somatic connectivity were introduced to the resonant model in §3.7. Travelling waves were observed when performing simulations in 1D and the resulting wavelength was determined analytically. Nevertheless, the continuation of the current time constant showed a minimal variation of the critical wave-number.

In summary, although mechanisms that associate the change of scale in the grid cells with model parameters (unrelated to their connectivity) were found, there is no experimental evidence to supports them. The next chapter explores more fully nonlinear models of the I_h currents in spiking networks and establishes that the distributions of such *rebound* current is a more viable mechanism for the local control of grid cell firing patterns.

4

AN INTEGRATE-AND-FIRE NETWORK MODEL FOR GRID CELL DYNAMICS

4.1 INTRODUCTION

In the previous chapter, a variety of firing rate network models was proposed in order to find a mechanism for the different scales of grid cell firing fields, based on their inhibitory connectivity and resonance properties. This chapter explores a different type of model inspired by experimental evidence that neural rebound spiking may provide a mechanism for differences in spacing of grid cell firing fields [51]. Thus, an integrate-and-fire network model with *postinhibitory rebound* (PIR) is proposed. The PIR modifies the intrinsic excitability of a neuron by evoking one or more spikes in response to the cessation of a prolonged hyperpolarising step or to a brief hyperpolarising stimulus [52]. When neurons are coupled with mutual inhibition, the PIR initiates a rhythmic activity pattern.

Some of the intrinsic properties of grid cells, like the theta rhythmic firing, have been addressed by the oscillatory interference models that use interference patterns generated by multiple membrane potential oscillations (MPOs) to explain grid formation [53]. This type of model takes into account the theta rhythmic firing of grid cells (5-12 Hz) and the fact that the spiking output increases in frequency as the animal approaches the centre of the grid cell firing field, known as *phase precession*, and is suggestive of a spike-timing code [54]. Moreover, there is experimental evidence that the frequency of intrinsic MPOs

decreases along the dorsal-ventral axis of MEC II [46]. Hence, these oscillations are hypothesised to play a role in grid field spacing. Other properties of grid cells, such as their inhibitory connectivity, are ubiquitous in MEC stellate cells, that are believed to represent the majority of the grid cell population [46]. Interestingly, the rapid membrane time constant and resonant behaviour[1] of stellate cells was studied before the discovery of grid cells. The resonant behaviour of stellate cells has been directly linked to a high density of hyperpolarisation-activated cyclic-nucleotide-gated (HCN) channels [55]. These channels generate the so-called I_h current that can cause a depolarising rebound[2] spike after a hyperpolarising current is injected. As mentioned in §3.7, there is experimental evidence that the time constant of both the fast and slow component of I_h is significantly faster for dorsal versus ventral stellate cells of the MEC II [50].

Some grid cell models have considered the intrinsic biophysical properties of stellate cells. For example: Navratilova *et al.* [56] incorporated post-spike dynamics that contribute to the theta precession, and the group headed by Hasselmo included HCN channels in their models [51, 57–60]. Hasselmo [51] proposed a spiking network model of MEC that can generate consistent firing fields despite the changes in direction[3] and running speed of the rat. This model shows that a difference of grid cell firing field spacing can be obtained by manipulating the speed of the rebound response of stellate cells. As mentioned previously, the rebound spiking is generated by I_h current dynamics; and therefore changing the parameters related to the I_h current in the model affects the firing field spacing. In the model proposed by Hasselmo [51], stellate cells are connected via inhibitory interneurons which receive input from the medial septum and head-direction cells. In [51], the running speed of the simulated rat is controlled by changing the frequency of the medial septum input or by increasing the inhibitory interneurons' input strength to stellate cells, whereas

1 Here, resonant behaviour refers to the increase in the likehood of firing when the neurons have inputs at a preferred resonant frequency.

2 In general, a rebound is a overshoot of a membrane potential above its resting level [52].

3 In the model proposed in [51], the simulated rat is constrained to a narrow rectangular environment, hence it can go only towards the left or the right ends of such an environment.

the direction is selectively controlled by head-direction input to inhibitory interneurons. The emergent stellate cell activity and firing fields are shown in Figure 4.1. Note that the activity of stellate cells (Fig. 4.1b) resembles a periodic travelling wave whose period relates to the spacing of the firing fields (Fig. 4.1c). Importantly, a change in the spacing of grid cell firing fields can be obtained by shifting only a single parameter influencing the intrisic rebound spiking and the resonant frequency of stellate cells. Additionally, this model exhibits phase precession in response to phasic medial septum input, theta cycle spiking and the loss of the spatial periodicity of grid cell firing fields upon a reduction of input from the medial septum. However, the study of computational spiking models is heavily reliant on computational approaches, and hence analytical insight is somewhat limited. Analytical insight is important because it allows predictions of the model behaviour under parameter variation and also determines which parameters control the spacing of grid cell firing fields.

This chapter proposes a new spiking network model (§4.2) motivated by the observations of Hasselmo [51]. The proposed model includes ideas from models of the I_h current (particularly from the model in [61]). In order to allow for analytical insight, only one population of stellate cells is modelled without (medial septum or head-direction) input. Computational simulations of the proposed model show periodic travelling waves with a period controlled by the dynamics of the I_h current. Hence, the focus of this chapter is to determine analytically the relationship between the model parameters and the travelling wave period. It is assumed that these periodic travelling waves resemble the activity of grid cells and that the temporal period of such travelling waves relates to the spacing between firing fields, as in the model proposed by Hasselmo [51] (see Fig. 4.1). But the specific relationship between the spacing of the firing fields of a simulated rat trajectory and the model parameters is not addressed in this work.

Here, a piecewise linear (PWL) function for HCN channel activation is proposed with the aim to describe analytically the travelling wave and maximum wavelength emerging from simulations. The travelling wave orbit is constructed

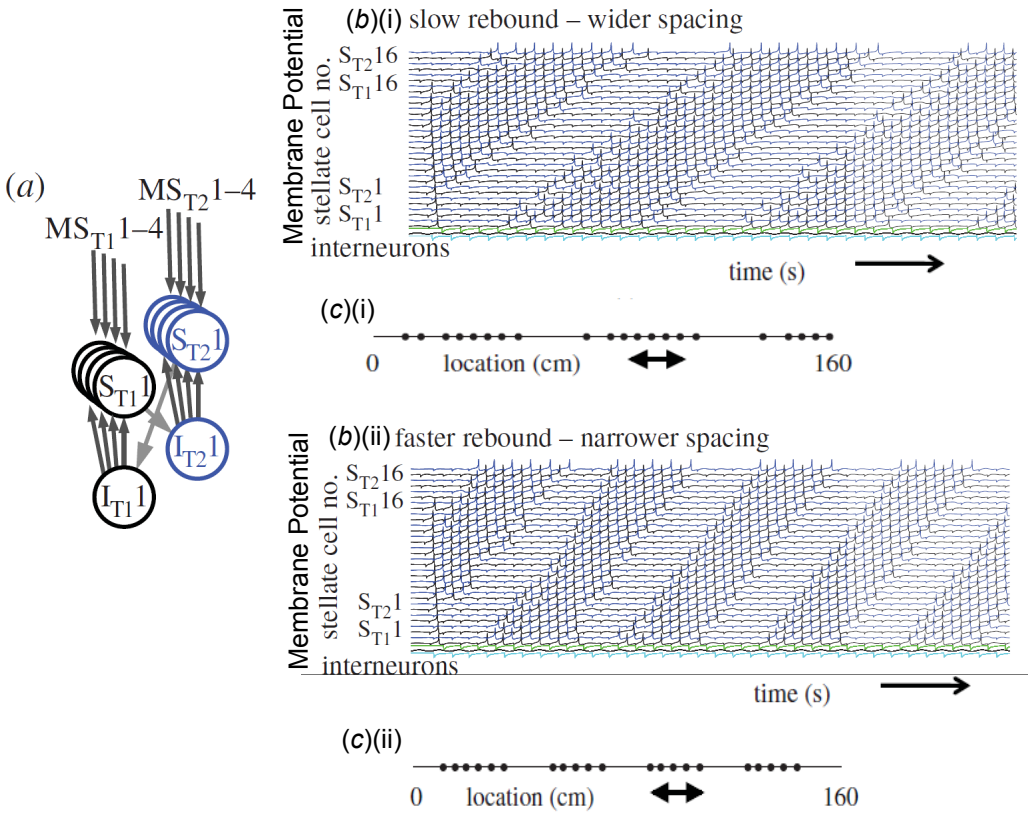


Figure 4.1. Model proposed by Hasselmo [51] using resonance and rebound spiking in medial entorhinal cortex. (a) Local circuit containing two populations (T1 and T2) of stellate cells (S), inhibitory interneurons (I) and medial septal input (MS). Numbers indicate the number of the cell within a population. Arrows indicate the connections between neurons. (b) Membrane potential of stellate cells in larger populations (16 stellate cells in each population) with (i) slow and (ii) fast rebound. Activity in population T1 is represented by black lines and in population T2 by blue lines. Note that the rebound spiking occurs in a subset of stellate cells and shifts progressively across the population. The bottom traces of these plots show the activity of both populations of interneurons (light blue and green). (c) The line denotes the trajectory of the simulated rat, going from one end to the other. Dots represent the position of the simulated rat where the firing events of a single stellate cell occurred (firing fields). The spacing between the firing fields is wider when the stellate cells have slow rebound. Figure adapted from [51].

in §4.3 and its stability is analysed in §4.4 using techniques from non-smooth dynamical systems. Furthermore, the stability analysis is validated in §4.5. Continuation of the parameters related to the I_h current is performed in §4.6.

4.2 THE MODEL

In this section a linear IF model is proposed instead of the Izhikevich style model (see [9] for this type of model) used by Hasselmo [51]. In the model, a spike is generated when the voltage exceeds a certain value V_{th}, after which it is reset to the value V_r, and the voltage evolution is described by (2.11). Additionally, the HCN channel dynamics are incorporated and a smooth inverse Top Hat synaptic connectivity in one spatial dimension is used (as in §3.6.5), based on experimental observations in [26], instead of the local circuit inhibitory and excitatory connectivity between different populations of stellate cells and interneurons proposed by Hasselmo. Further, a refractory process is included, such that the voltage is held at V_r after a spike event for a duration of τ_R. This refractory process disallows the generation of a burst of spikes; and therefore, simplifies the model behaviour which in turn facilitates the tractability of the analysis. Numerical simulations are performed in order to investigate if the proposed model is able to exhibit travelling waves, as those described in the model by Hasselmo (but without the bursting behaviour).

From (2.43) the synaptic current for this model is rewritten as

$$I_{syn}(x,t) = g_{syn}\psi(x,t), \tag{4.1}$$

where g_{syn} is the overall strength of synaptic conductance and

$$\psi(x,t) = \sum_{m\in\mathbb{Z}} \int_{-\infty}^{\infty} \mathrm{d}y W(x-y)\eta(t-T^m(y)), \qquad x \in \mathbb{R}, \qquad t > 0. \tag{4.2}$$

Here W is assumed to be symmetric, thus $W(x) = W(|x|)$ and η is a synaptic filter given by (2.22). Particularly, the following smooth inverse Top Hat connectivity is used:

$$W(x) = \frac{W_0}{2}\left(\tanh\left(\varsigma\left(R-|x|\right)\right) + \tanh\left(\varsigma\left(R+|x|\right)\right)\right). \tag{4.3}$$

Figure 4.2 shows this connectivity function. Note that $T^m(x)$ in (4.2) represents the time of arrival of the mth presynaptic spike at position x generated according to

$$T^m(x) = \inf\left\{t \mid V(x,t) \geq V_{th}; t \geq T^{m-1}(x) + \tau_R\right\}, \qquad m \in \mathbb{Z}. \tag{4.4}$$

The I_h current is initially modelled as in §2.2.1 with a single inactivation gate, so that

$$I_h(x,t) = g_h n_h(x,t)(V_h - V(x,t)), \tag{4.5}$$

where

$$\tau_h(V)\frac{\partial n_h}{\partial t} = n_{h,\infty}(V) - n_h. \tag{4.6}$$

Here $n_{h,\infty}$ is a sigmoid with

$$n_{h,\infty}(V) = \left[1 + \exp((V - V_{1/2})/k)\right]^{-1}, \tag{4.7}$$

and fits to experimental data give $V_{1/2} \approx -10$ mV (with respect to rest) and $k \approx 10$ [62,63] (Fig. 4.3c). The time constant τ_h can vary from tens to hundreds of ms [50]. For simplicity, the voltage dependence of τ_h is neglected and it is viewed as a fixed constant model parameter.

Numerical simulations for this model are performed as described in Appendix F. Travelling waves were observed after applying a hyperpolarising current to a set of neurons for a few milliseconds (Fig. 4.3a).

Bearing in mind that the goal of this chapter is to mathematically analyse the emerging waves, the model is reduced by:

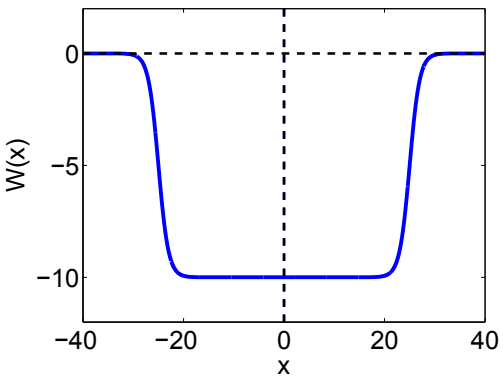


Figure 4.2. Connectivity function W in (4.3) with $\varsigma = 0.5$, $R = 25$ and $W_0 = -10$.

- Absorbing V_h in g_h since V_h is typically larger than V, which suggests $V_h - V \simeq V_h$.
- Translating the rest state from $V = -60$ mV to $V = 0$ mV.
- Approximating $n_{h,\infty}$ with a PWL function (Fig. 4.3c).

Importantly, numerical simulations show similar results after these changes (Fig. 4.3b). Therefore, the model describing the voltage $V = V(x,t)$ for the neuron at position $x \in \mathbb{R}$ and time t adopted in this chapter is given by:

$$C\frac{\partial V(x,t)}{\partial t} = -g_l V(x,t) + I_h(x,t) + I_{syn}(x,t). \tag{4.8}$$

Here $I_h(x,t) = g_h n_h(x,t)$, with $n_h(x,t)$ governed by (4.6) (in which $\tau_h(V)$ is constant for simplicity) and the PWL approximation of $n_{h,\infty}(V)$ given by

$$n_{h,\infty}(V) = \begin{cases} 1, & V \le V_- \\ \frac{1}{2} - \frac{V - V_{1/2}}{4k}, & V_- < V \le V_+, \\ 0, & V \ge V_+ \end{cases} \qquad V_\pm = V_{1/2} \pm 2k. \tag{4.9}$$

Figure 4.3b shows an example of the emerging travelling waves in the simplified system (see Appendix F for simulation details). The spiking activity is initiated after applying a hyperpolarising current to a set to neurons. The period and speed of the travelling wave can be easily calculated using a raster plot such as that shown in Figure 4.3d. Note that there are some boundary effects. These boundary effects arise because the network is aperiodic, i.e., the opposite edges of the line, where neurons are arranged, are disconnected. Hence, neurons at the edges connect with fewer neurons; and therefore, receive less inhibitory input. Figure 4.3e shows that n_h activates when the neuron is hyperpolarised, as expected for HCN channels. Importantly, the emergence of travelling waves is robust to the selection of duration and strength of the initial hyperpolarising current, as well as the set of neurons where the current is applied. Indeed, it emerges even from initial hyperpolarisation of a randomly chosen set of neurons (data omitted). Such variations in initialisation (independent of model parameter choice) do influence the period and speed of the resulting travelling

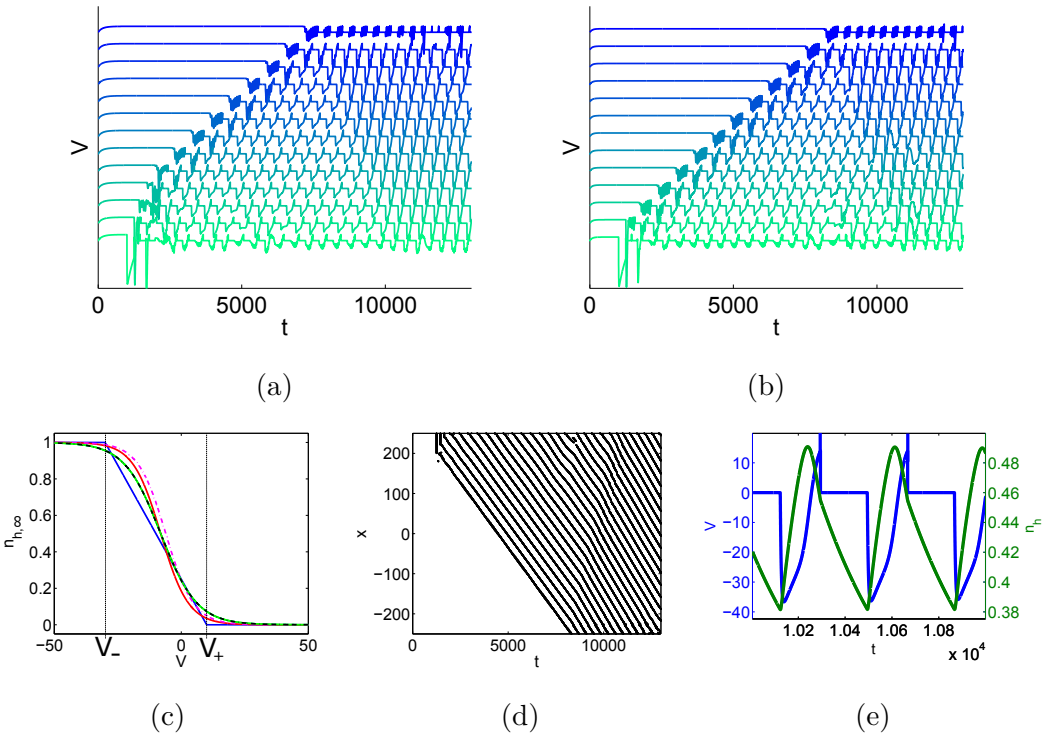


Figure 4.3. (a) Simulations of a spiking neural field model (4.8) with a sigmoid function for the I_h activation (4.7). Here, the voltage traces at a subset of locations throughout the system are shown as a function of time. (b) Same as (a) but with a piecewise linear (PWL) function for I_h activation (4.9) (blue line in (c)). (c) Activation function for n_h from experimental data [63] (after translating the steady state from 60mV to 0mV) for fast dorsal (green line) and ventral (red line) activation; and slow dorsal (black line) and ventral (magenta line) activation. Blue line: PWL fit of $n_{h,\infty}$ given by (4.9) with $V_{1/2} = -10$ mV and $k = 10$. (d): Raster plot of the complete system in (b), each dot represents a spike at a certain location and time. (e) Evolution of the voltage (blue line) and n_h (green line) in time for a neuron (taken from the simulation in (b)). Paramenters: $C = 1\mu$ Fcm^{-2}, $\tau_h = 400$ ms, $V_h = 40$ mV, $g_l = 0.25$ mS/cm^{-2}, $g_h = 1$ mS/cm^{-2}, $\tau_R = 200$ ms, $V_{th} = 14$ mV, $V_r = 0$ mV. For the synaptic input $g_{syn} = 15$ mS/cm^{-2}, $W_0 = -10$, $R = 25$, $\varsigma = 0.5$ and $\alpha^{-1} = 20$ms. Here, an initial hyperpolarisation current of -30mV was applied to a set of 500 neurons (at the top in (d)) from $t = 1000$ ms to $t = 1250$ ms. The system consists of 5000 neurons. Numerical methods for simulation are provided in Appendix F.

wave. The speeds can be either positive or negative, and for simplicity only the positive ones are considered in the following analysis (the negative wave speed can be obtained by symmetry).

4.3 TRAVELLING WAVE FRAMEWORK

As shown in Figure 4.3 the system settles to a travelling wave with a well defined speed and period. In order to understand how I_h controls the emergent scale of such waves, the vector $X = (V, n_h) \in \mathbb{R}^2$ is introduced and the model is written in terms of the PWL evolution equation that governs the system behaviour between one spiking event and the next:

$$\frac{\partial X(x,t)}{\partial t} = A(x,t)X(x,t) + \Psi(x,t), \qquad T^{m-1}(x) \leq t < T^m(x), \tag{4.10}$$

where A and Ψ are defined depending on V, or whether the system is in the refractory state. The dependence on V dictates the value of the proposed PWL description of n_{h_∞} (4.6). In this way, the evolution of the system is described by (4.10) with the appropriate selection of A and Ψ depending on the state of the system. In detail:

$$A(x,t) = \begin{cases} A_R = \begin{bmatrix} 0 & 0 \\ 0 & -1/\tau_h \end{bmatrix}, & T^{m-1}(x) \leq t < T^{m-1}(x) + \tau_R \\ A_\mu, & T^{m-1}(x) + \tau_R \leq t < T^m(x) \end{cases}, \tag{4.11}$$

with

$$A_\mu = \begin{cases} A_0 = \begin{bmatrix} -1/\tau & g_h g_l^{-1}/\tau \\ -1/(4k\tau_h) & -1/\tau_h \end{bmatrix}, & V_- < V < V_+ \\ A_- = \begin{bmatrix} -1/\tau & g_h g_l^{-1}/\tau \\ 0 & -1/\tau_h \end{bmatrix}, & V \leq V_- \\ A_+ = A_-, & V \geq V_+ \end{cases}. \tag{4.12}$$

In (4.12) the ordering $V_- < V_r < V_+ < V_{th}$ has been assumed and $\tau = C/g_l$ has been introduced. Similarly Ψ is defined by

$$\Psi = \begin{cases} \Psi_R = \begin{bmatrix} 0 \\ \frac{1/2-(V_r-V_{1/2})/(4k)}{\tau_h} \end{bmatrix}, & T^{m-1}(x) \le t < T^{m-1}(x) + \tau_R \\ \Psi_\mu = \frac{g_{syn}g_l^{-1}}{\tau}\psi \begin{bmatrix} 1 \\ 0 \end{bmatrix} + b_\mu, & T^{m-1}(x) + \tau_R \le t < T^m(x) \end{cases}, \tag{4.13}$$

where

$$b_\mu = \begin{cases} b_0 = \begin{bmatrix} 0 \\ \frac{1/2-(V_r-V_{1/2})/(4k)}{\tau_h} \end{bmatrix}, & V_- < V < V_+ \\ b_- = \begin{bmatrix} 0 \\ 1/\tau_h \end{bmatrix}, & V \le V_- \\ b_+ = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, & V \ge V_+ \end{cases}. \tag{4.14}$$

The above notation allows indication of the state of the system, whether it is refractory (labelled by R) or if not refractory and has a voltage in the range (V_-, V_+) (labelled by 0), $(-\infty, V_-]$ (labelled by $-$) or $[V_+, \infty)$ (labelled by $+$), see Figure 4.3c.

In order to describe the emerging travelling waves, solutions of (4.10) of the form $\overline{X}(\xi, t)$ are investigated, where $\xi = t - x/c$ and c is the (constant) wave speed. Therefore $\overline{X}(\xi, t)$ is given by

$$\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial \xi}\right)\overline{X}(\xi, t) = A\overline{X}(\xi, t) + \overline{\Psi}(\xi, t). \tag{4.15}$$

The system eventually settles to a stationary travelling wave, as shown in Figure 4.3. This wave is denoted by $\overline{X}(\xi, t) = Q(\xi) = (V(\xi), n_h(\xi))$ and satisfies the travelling wave equation:

$$\frac{\mathrm{d}Q(\xi)}{\mathrm{d}\xi} = AQ(\xi) + \Psi(\xi). \tag{4.16}$$

A firing event in the periodic wave is described by $T^m(x) = x/c - m\Delta$, where Δ is the period of the wave such that $Q(\xi + m\Delta) = Q(\xi)$ and $m \in \mathbb{Z}$. Substituting in (4.2) gives

$$\psi(\xi) = \psi(x,t)|_{T^m(x)=x/c-m\Delta} = c \sum_{m\in\mathbb{Z}} \int_0^\infty \mathrm{d}s\eta(s)W\left(c(s-\xi) + cm\Delta\right). \tag{4.17}$$

Hence $\Psi(\xi)$ in (4.16) is obtained from (4.13) under the replacement of $\psi(x,t)$ by $\psi(\xi)$. Since $\psi(\xi)$ is Δ-periodic it can be expressed in terms of a Fourier series as

$$\psi(\xi) = \sum_{p\in\mathbb{Z}} \psi_p \mathrm{e}^{-2\pi i p\xi/\Delta}, \qquad \psi_p = \frac{1}{\Delta}\widetilde{W}\left(\frac{2\pi p}{c\Delta}\right)\widetilde{\eta}\left(-\frac{2\pi p}{\Delta}\right), \tag{4.18}$$

where $\widetilde{\ }$ denotes the Fourier transform, given by (A.3), of the connectivity matrix W in (4.3) and the α-function in (2.22), namely

$$\widetilde{W}(k) = W_0\frac{\pi}{\varsigma}\frac{\sin(kR)}{\sinh\left(\pi k/(2\varsigma)\right)}, \qquad \widetilde{\eta}(k) = \frac{\alpha^2}{(\alpha+ik)^2}. \tag{4.19}$$

Here, the result that $c\Delta\sum_m \mathrm{e}^{ikcm\Delta} = 2\pi\sum_p \delta(k - 2\pi p/(c\Delta))$ was exploited, where $\sum_p \delta(k - 2\pi p/(c\Delta))$ represents an infinite sequence of unit impulses spaced at $2\pi p/(c\Delta)$ with Fourier transform $(c\Delta\sum_m \mathrm{e}^{ikcm\Delta})/(2\pi)$ [64]. Note that both expressions in (4.19) decay as a function of k, and therefore the infinite sum in (4.18) can be naturally truncated.

The formal solution to (4.16) is constructed using variation of parameters:

$$Q(\xi) = G(\xi,\xi_0)Q(\xi_0) + \int_{\xi_0}^{\xi} \mathrm{d}\xi' G(\xi,\xi')\Psi(\xi'), \tag{4.20}$$

where G is a matrix exponential given by

$$G(\xi,\xi') = \mathcal{T}\left\{\exp\left(\int_{\xi'}^{\xi} \mathrm{d}sA(s)\right)\right\}, \tag{4.21}$$

with $\mathcal{T}$ a time-ordering operator $\mathcal{T}\{A(t)A(s)\} = H(t-s)A(t)A(s) + H(s-t)A(s)A(t)$, in which $H(t)$ is the Heaviside step function defined in (2.23). Note that in this case A is piecewise constant, thereby simplifying the evaluation of G. Accordingly, the solution is broken into parts distinguished by the label $\mu \in \{R, 0, -, +\}$, such that the solution of (4.16) is explicitly given by (4.20)

with $G(\xi,\xi') = G(\xi-\xi')$ and $G = G_\mu$, where $G_\mu(\xi) = \exp(A_\mu \xi)$. The shape of the periodic travelling wave can be constructed in a self-consistent manner by patching together solutions Q_μ for each domain, as shown below.

4.3.1 *Travelling Wave Orbit Construction*

Properties of matrix exponential are used in order generate a useful representation of (4.20) for each domain. A general matrix exponential can be written in the form $\mathrm{e}^{At} = P\mathrm{e}^{\Lambda}P^{-1}$, where $\Lambda = \mathrm{diag}(\lambda^+,\lambda^-)$ comprises the eigenvalues of A given explicitly by

$$\lambda^{\pm} = \frac{1}{2}\left(\mathrm{Tr}(A) \pm \sqrt{(\mathrm{Tr}(A))^2 - 4\mathrm{det}(A)}\right), \tag{4.22}$$

and the associated eigenvectors are given by $q^{\pm} = (1,(\lambda^{\pm}-A_{11})/A_{12})^T$, so that $P = [q^+,q^-]$. By using this representation for matrix exponentials in conjunction with (4.18), a domain specific trajectory for $\mu \in \{0,+,-\}$ can be written as

$$\begin{aligned} Q_\mu(\xi) &= G_\mu(\xi-\xi_0)Q_\mu(\xi_0) + A_\mu^{-1}\left[G_\mu(\xi-\xi_0) - I_2\right] b_\mu \\ &+ \frac{g_{syn}g_l^{-1}}{\tau}\sum_{p\in\mathbb{Z}} \psi_p P_\mu \mathrm{diag}(Z_\mu^+(\xi,\xi_0), Z_\mu^-(\xi,\xi_0)) P_\mu^{-1} \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \end{aligned} \tag{4.23}$$

where

$$Z_\mu^{\pm}(\xi,\xi_0) = \frac{\mathrm{e}^{\lambda_\mu^{\pm}(\xi-\xi_0)}\mathrm{e}^{-2\pi i p\xi_0/\Delta} - \mathrm{e}^{-2\pi i p\xi/\Delta}}{\lambda_\mu^{\pm} + 2\pi i p\Delta}. \tag{4.24}$$

Note that if $\mu = R$ the system is refractory with $V(\xi)$ clamped at $V = V_r$, and the corresponding A_R in (4.11) is singular (its determinant is equal to zero). In this case, only the evolution of the gating variable n_h (4.6) is considered, and using the PWL activation function (4.9) gives

$$n_h(\xi) = n_h(\xi_0)\mathrm{e}^{-(\xi-\xi_0)/\tau_h} + \left(\frac{1}{2} - \frac{V_r - V_{1/2}}{4k}\right)\left[1 - \mathrm{e}^{-(\xi-\xi_0)/\tau_h}\right], \tag{4.25}$$

instead of (4.23).

Figure 4.4 shows a periodic travelling wave in the (V,n_h) phase-plane. This figure was obtained by plotting the solution of the system (4.23) for the different

system domains (except when $\mu = R$ for which (4.25) is used and V is clamped at $V = V_r$). The translational invariance of the system allows the choice of any point as the travelling wave origin. For simplicity $\xi = 0$ is selected, corresponding to the system immediately after firing.

- From $\xi = 0$ to $\xi = \tau_R$ the system remains clamped at V_r with n_h evolving according to (4.25) and $\xi_0 = 0$ (blue line).

- Then it evolves according to (4.23) with $\mu = 0$ and initial data determined by $Q_R(\xi_0) = (V_r, n_h(\tau_R))$, until $V(\xi)$ reaches $V_\pm$, after which the domain changes to $\mu = -$ or $\mu = +$ in (4.23), and appropriate initial data is selected depending on the value of V achieved first. Although some of the travelling wave orbits emerging from the performed simulations are restricted to $V > V_-$, for completeness it is assumed that after $\mu = 0$ the system switches to $\mu = -$ (cyan line in Fig. 4.4).

- The system stays in the state where $V < V_-$ until it returns to $V = V_-$ (green line in Fig. 4.4). Here the initial data is determined by $Q_-(\xi_0) = Q_0(\xi_1 + \tau_R)$.

- Then the system evolves according to (4.23) with $\mu = 0$ and initial conditions given by $Q_0(\xi_0) = Q_-(\xi_2 + \xi_1 + \tau_R)$ (yellow line in Fig. 4.4).

- Finally, when $V = V_+$, $\mu = +$ in (4.23) with $Q_+(\xi_0) = Q_0(\xi_3 + \xi_2 + \xi_1 + \tau_R)$ until $V(\xi) = V_{th}$ (orange line in Fig. 4.4), thus closing the orbit.

Here the time of flight for the trajectory is denoted by ξ_1 for $V_r \geq V > V_-$ (between release from refractoriness and reaching $V = V_-$), by ξ_2 for $V_- \geq V$, by ξ_3 for $V_- \leq V < V_+$, and by ξ_4 for $V_+ \leq V < V_{th}$. The period of the orbit is given by $\Delta = \tau_R + \xi_1 + \xi_2 + \xi_3 + \xi_4$. Note that the wave orbit in Figure 4.4 is discontinuous because of the voltage reset after one period. The theoretical orbits constructed in this way are in excellent agreement with numerical simulations, as shown in Figure 4.4.

In the above construction there are six unknowns $(n_h(0), c, \xi_1, \xi_2, \xi_3, \Delta)$ related by five nonlinear algebraic equations

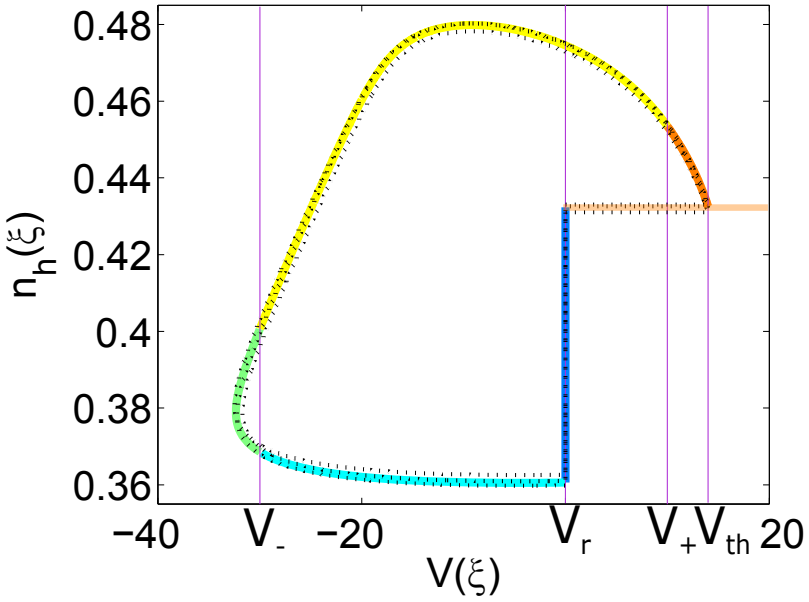


Figure 4.4. Orbit of a travelling wave. The travelling wave starts at $\xi = 0$ where V is clamped at V_r, hence n_h evolves according to (4.25) until $\xi = \tau_R$ and the system is released from the refractory period (blue line). Then it evolves according to (4.23) with $\mu = 0$ until $V(\tau_R + \xi_1) = V_-$ (cyan line) when it switches to $\mu = -$ until $V(\tau_R + \xi_1 + \xi_2) = V_-$ (green line). Afterwards, it goes back to $\mu = 0$ (yellow line) before switching to $\mu = +$ when $V(\tau_R + \xi_1 + \xi_2 + \xi_3) = V_+$ (orange line). The orbit ends when $V = V_{th}$ and V is reset (pale orange). Black dotted line: activity in the (V, n_h) plane of a neuron at position $x = 0$ when the simulation in Figure 4.3b settles to a travelling wave. Purple lines: values of V denoting the different domains of the system (μ in (4.23)). Parameters as in Figure 4.3. Here $\Delta = 393.53$, $c = 0.0779$, $\xi_1 = 8.2622$, $\xi_2 = 20.8468$, $\xi_3 = 145.3176$, $n_h(0) = 0.4323$.

1. $V(\Delta) = V_{th}$ (firing condition),
2. $n_h(\Delta) = n_h(0)$ (periodicity condition),
3. $V(\tau_R + \xi_1) = V_-$ (switching condition),
4. $V(\tau_R + \xi_1 + \xi_2) = V_-$ (switching condition),
5. $V(\tau_R + \xi_1 + \xi_2 + \xi_3) = V_+$ (switching condition),

whose simultaneous solution determines the dispersion relationship for the wave speed as a function of the period $c = c(\Delta)$. Figure 4.5 shows a dispersion curve constructed in this way for positive speeds. The corresponding curve for negative speeds is symmetric with respect to the Δ-axis (not shown). Interestingly, a wide range of allowed wavelengths can co-exist (with different speeds) as seen in simulations with various choices of initial hyperpolarisation strength and duration. Note that the constructed orbits change along the dispersion curve, the ones corresponding to small periods have $V < V_-$, whereas orbits with larger periods have $V > V_-$. However, the periods of orbits found in simulations are restricted to small values on the dispersion curve. Therefore, the aim of the rest of this chapter is to investigate the maximum period to which the travelling waves can settle. In order to achieve this, an analysis of wave stability under parameter variation is needed.

4.4 WAVE STABILITY

The model proposed in this chapter poses potential mathematical difficulties for standard approaches for linear stability analysis, such as that employed in Chapter 3, since:

- The perturbations of the state variables also affect the times of firing.
- The model switches when $V = V_\pm$, and hence is non-smooth.
- There is a discontinuity in the model when it resets after firing.

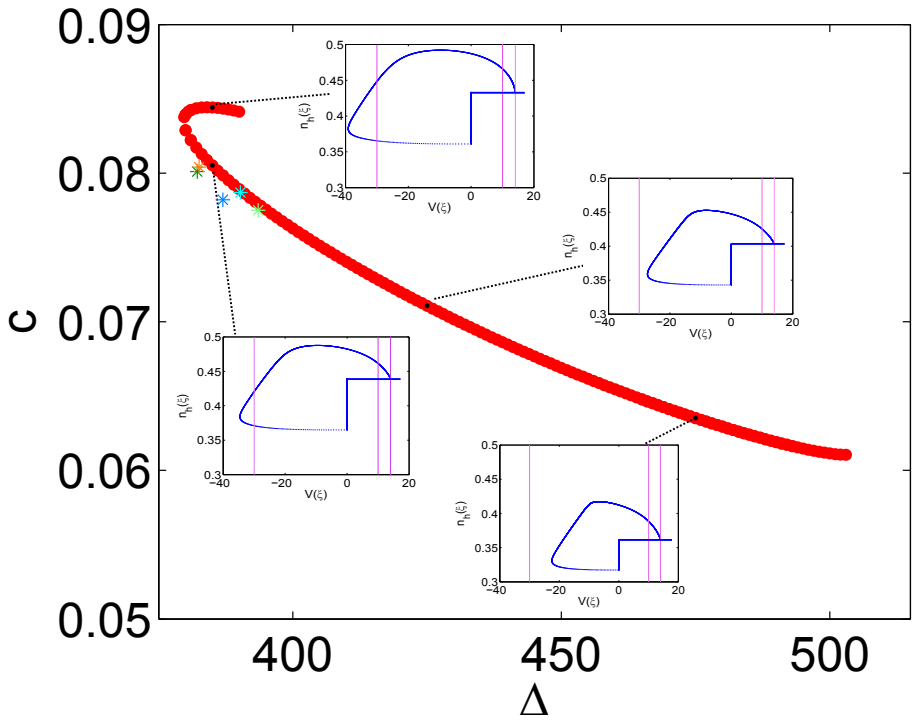


Figure 4.5. Dispersion curve $c = c(\Delta)$ for a periodic travelling wave (red dots) in the system (4.16). Asterisks represent values of the emerging travelling waves in simulations initiated with different strength and duration of an initial hyperpolarising current, applied to different subset of neurons. Light green corresponds to the simulation in Figure 4.4. Inset: travelling wave orbits constructed theoretically for different periods. For small periods and on the upper branch, constructed orbits visit the domain where $V < V_-$. Parameters as in Figure 4.3.

- After firing the voltage is clamped for a period of time τ_R.

Therefore, here the stability for a spiking neural field model with refractoriness is determined by constructing an Evans function. Such function was first used to determine the stability of action potentials in the Hodgkin-Huxley model [65], and more recently extended to certain classes of neural firing field models [66].

To facilitate the analysis, the spike-train that determines the synaptic drive in (4.2) is exposed by writing

$$\psi(x,t) = \sum_{m\in\mathbb{Z}} \int_{-\infty}^{\infty} \mathrm{d}y W(|x-y|) \int_{-\infty}^{t} \mathrm{d}s \eta(t-s)\delta(s-T^m(y)). \tag{4.26}$$

An alternative expression can be obtained for the spike events $\delta(t-T^m(x))$ by noting that for a fixed x:

$$\delta(t-T^m(x)) = |V_t(x,T^m(x))|\,\delta(V(x,t)-V_{th}), \tag{4.27}$$

where the subscript t denotes partial differentiation with respect to t. The following property of Dirac δ functions has been used :

$$\delta(\gamma(x)) = \sum_n \left|\gamma\,'(x_n)\right|^{-1} \delta(x-x_n), \qquad \gamma(x_n)=0, \qquad \gamma\,'(x_n)\neq 0. \tag{4.28}$$

Substituting (4.27) into (4.26) gives

$$\psi(x,t) = \sum_{m\in\mathbb{Z}} \int_{-\infty}^{\infty} \mathrm{d}y W(|x-y|) \int_{-\infty}^{t} \mathrm{d}s \eta(t-s)\,|V_t(y,T^m(y))|\,\delta(V(y,s)-V_{th}). \tag{4.29}$$

A co-moving frame is considered in order to describe how perturbations of a periodic travelling wave evolve in time. In this co-moving frame the solutions are of the form $V(x,t) = \overline{V}(\xi,t)$ with $\xi = t - x/c$; and the firing events are defined as functions $\xi_m(t)$ according to the threshold condition $\overline{V}(\xi_m(t),t) = V_{th}$. Using (4.28) yields

$$\delta(\xi-\xi_m(t)) = \left|\overline{V}_\xi(\xi_m(t),t)\right|\delta(\overline{V}(\xi,t)-V_{th}). \tag{4.30}$$

Substitution of (4.30) in (4.29), together with the assumption that $\overline{V}_\xi \simeq V_t$ close to a periodic orbit, gives

$$\begin{aligned}\psi(x,t) &= \sum_{m\in\mathbb{Z}} \int_{-\infty}^{\infty} \mathrm{d}y W(|x-y|) \int_{-\infty}^{t} \mathrm{d}s \eta(t-s)\delta(s-y/c-\xi_m(s)) \\ &= c\sum_{m\in\mathbb{Z}} \int_0^{\infty} \mathrm{d}s \eta(s) W\left(|c(s-\xi)+c\xi_m(t-s)|\right) \equiv \overline{\psi}(\xi,t).\end{aligned} \tag{4.31}$$

For a periodic wave $\xi_m(t) = m\Delta$; therefore $\overline{\psi}(\xi, t)$ is independent of t and (4.31) is equal to (4.17).

To analyse stability, small perturbations to this trajectory are considered. Note that the firing times of the perturbed and unperturbed trajectories will differ by a small amount defined here by $\delta\xi_m(t)$, such that a perturbed firing time $\check{\xi}_m(t)$ is given by

$$\check{\xi}_m(t) = m\Delta + \delta\xi_m(t). \tag{4.32}$$

Substitution in (4.31) gives the corresponding perturbation of $\overline{\psi}(\xi, t)$ as $\check{\overline{\psi}}(\xi, t) = \psi(\xi) + \delta\overline{\psi}(\xi, t)$ where

$$\delta\overline{\psi}(\xi, t) = c^2 \sum_{m\in\mathbb{Z}} \int_0^\infty \mathrm{d}s\eta(s)W'\left(|c(s-\xi) + cm\Delta|\right)\delta\xi_m(t-s). \tag{4.33}$$

4.4.1 *Perturbed Firing Event Characterisation*

In order to describe how the perturbation behaves when the system passes through the values V_+, V_-, and V_{th}, an indicator function is introduced. This function is denoted by $h(\overline{X}(\xi, t); \vartheta) = \overline{V}(\xi, t) - V_\vartheta$, where $\vartheta \in \{+, -, \mathrm{th}\}$, and allows one to define the travelling wave coordinate values at which switching events occur as $h(\overline{X}(\xi, t); \vartheta) = 0$. As in §4.3 $\overline{X}(\xi, t) = Q(\xi)$ denotes an unperturbed travelling wave trajectory, whereas a perturbed one is given by

$$\check{\overline{X}}(\xi, t) = Q(\xi) + \delta\overline{X}(\xi, t), \tag{4.34}$$

with $\delta\overline{X}$ small. Let $\xi = \xi_s(t)$ denote when an unperturbed trajectory passes through the switching manifold, and

$$\check{\xi}_s(t) = \xi_s(t) + \delta\xi_s(t), \tag{4.35}$$

for a perturbed trajectory. Taylor expanding the indicator function gives:

$$\begin{aligned} h(\check{\overline{X}}(\check{\xi}_s, t)) &= h(\check{\overline{X}}(\xi_s + \delta\xi_s, t)) = h(Q(\xi_s + \delta\xi_s) + \delta\overline{X}(\xi_s + \delta\xi_s, t)) \\ &\simeq h(Q(\xi_s) + Q_\xi(\xi_s^-)\delta\xi_s) + \nabla_Q h(Q(\xi_s + \delta\xi_s)) \cdot \delta\overline{X}(\xi_s + \delta\xi_s, t) \\ &\simeq h(Q(\xi_s)) + \nabla_Q h(Q(\xi_s)) \cdot Q_\xi(\xi_s^-)\delta\xi_s + \nabla_Q h(Q(\xi_s)) \cdot \delta\overline{X}(\xi_s, t). \end{aligned} \tag{4.36}$$

Here

$$\overline{X}(\xi_s^{\pm}, t) = \lim_{\epsilon \searrow 0} \overline{X}(\xi_s \pm \epsilon, t), \tag{4.37}$$

ensures that the partial derivative in ξ is well defined. By definition $h(Q(\xi_s)) = 0 = h(\check{\overline{X}}(\check{\xi}_s, t))$; substitution in (4.36) gives

$$\nabla_Q h(Q(\xi_s)) \cdot \left[\delta\overline{X}(\xi_s, t) + Q_\xi(\xi_s^-)\delta\xi_s\right] = 0, \tag{4.38}$$

where $\nabla_Q h(Q; \vartheta) = (\partial_V, \partial_{n_h})(V - V_\vartheta) = (1, 0)$. Rearranging terms yields

$$\delta\xi_s(t) = -\frac{\delta\overline{V}(\xi_s, t)}{V_\xi(\xi_s^-)}. \tag{4.39}$$

The above equation defines the perturbed switching event in terms of the difference between the perturbed and unperturbed trajectories. In the case of a firing event, (4.39) transforms to:

$$\delta\xi_m(t) = -\frac{\delta\overline{V}(m\Delta, t)}{V_\xi(m\Delta^-)}. \tag{4.40}$$

Hereinafter, it is assumed that the perturbed trajectory is separable in the sense that it can be decomposed into a periodic function in the travelling wave coordinate and an exponential in time. Thus

$$\delta\overline{X}(\xi, t) = \delta X(\xi)\mathrm{e}^{\lambda t}, \qquad \delta X(\xi) = \delta X(\xi + m\Delta), \qquad m \in \mathbb{Z}. \tag{4.41}$$

Finally, substituting (4.40) with (4.41) in (4.33) results in $\delta\overline{\psi}(\xi, t) = \delta\psi(\xi; \lambda)\mathrm{e}^{\lambda t}$, with $\delta\psi(\xi; \lambda) = \delta V(0) f(\xi; \lambda)$, and

$$f(\xi; \lambda) = -\frac{c^2}{V_\xi(0)} \sum_{m \in \mathbb{Z}} \int_0^\infty \mathrm{d}s \eta(s) W'\left(|c(s - \xi) + cm\Delta|\right) \mathrm{e}^{-\lambda s}. \tag{4.42}$$

4.4.2 *Evolving Perturbations through Switching Manifolds*

Now that the effect of perturbations on the synaptic component of the proposed model through the firing events has been described, it is possible to analyse how such perturbations evolve. Linearisation around the travelling wave by setting $\overline{X}(\xi, t) = Q(\xi) + \delta\overline{X}(\xi, t)$ in (4.15) with perturbations of the form (4.41) gives:

$$\frac{\mathrm{d}}{\mathrm{d}\xi}\delta X(\xi) = A(\xi; \lambda)\delta X(\xi) + \delta\Psi(\xi; \lambda), \tag{4.43}$$

where

$$\delta\Psi(\xi;\lambda) = \frac{g_{syn}g_l^{-1}}{\tau}\delta\psi(\xi;\lambda)\begin{bmatrix}1\\0\end{bmatrix}, \tag{4.44}$$

and $A(\xi;\lambda) = A(Q(\xi)) - \lambda I_2$, with $A(Q(\xi))$ defined in (4.11) and depending on the position along the periodic orbit; I_2 denotes the 2×2 identity matrix. As in §4.3, the solution for the orbits with $V > V_-$ can be written using a variation of parameters formula, matrix exponentials and (4.42) in the following way

$$\delta X(\xi) = \begin{cases} G_R(\xi;\lambda)\delta X(0), & 0 \le \xi < \tau_R \\ G_0(\xi-\tau_R;\lambda)\delta X(\tau_R) & \\ \quad + \int_{\tau_R}^{\xi} \mathrm{d}\xi' G_0(\xi-\xi';\lambda) J f(\xi';\lambda)\delta X(0), & \tau_R \le \xi < \tau_R+\xi_1 \\ G_+(\xi-(\tau_R+\xi_1);\lambda)\delta X(\tau_R+\xi_1) & \\ \quad + \int_{\tau_R+\xi_1}^{\xi} \mathrm{d}\xi' G_+(\xi-\xi';\lambda) J f(\xi';\lambda)\delta X(0), & \tau_R+\xi_1 \le \xi < \Delta \end{cases}. \tag{4.45}$$

Here $G_\mu(\xi;\lambda) = \exp\left(\left[A_\mu - \lambda I_2\right]\xi\right)$ and

$$J = \frac{g_{syn}g_l^{-1}}{\tau}\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}. \tag{4.46}$$

This solution can be extended for the orbits with $V < V_-$. Note that there is a jump in the Jacobian at $V = V_\pm$ (switching manifolds), $V = V_{th}$ (firing threshold) and when the system is released from the refractory period. The evolution of perturbations through these points needs careful handling. First, the discontinuity in the Jacobian due to the firing event is addressed following Coombes *et al.* [67]. Note that after a firing event, there is a jump in the voltage from V_{th} to V_r, whereas the value of n_h is continuous (see Fig. 4.6). Figure 4.7 shows that the perturbed trajectory can reach the threshold before or after the unperturbed trajectory. To describe how the perturbed trajectory resets after reaching the threshold, a vector function that governs the discontinuity at the firing event is introduced as $g(\overline{X}(\xi,t)) = (V_r, \overline{n}_h(\xi,t))^T$. For the unperturbed trajectory $Q(\xi_m^+) = g(Q(\xi_m^-))$ (with the notation in (4.37)), meaning that after firing the voltage is clamped to V_r and n_h evolves freely. Similarly, for the perturbed trajectory

$$\breve{\overline{X}}(\breve{\xi}_m^+, t) = g(\breve{\overline{X}}(\breve{\xi}_m^-, t)). \tag{4.47}$$

Taylor expanding (4.47) and using (4.41) gives

$$\breve{\overline{X}}(\breve{\xi}_m^+, t) \simeq Q(\xi_m^+) + Q_\xi(\xi_m^+)\delta\xi_m + \delta X(\xi_m^+)\mathrm{e}^{\lambda t}, \tag{4.48}$$

and

$$\begin{aligned} g(\breve{\overline{X}}(\breve{\xi}_m^-, t)) &\simeq g(Q(\xi_m^-)) + \nabla_Q^T g^T(Q(\xi_m^-))\delta\xi_m Q(\xi_m^-) \\ &+ \nabla_Q^T g^T(Q(\xi_m^-))\delta X(\xi_m^-)\mathrm{e}^{\lambda t}, \end{aligned} \tag{4.49}$$

with

$$\nabla_Q^T g^T = \begin{pmatrix} \partial V \\ \partial n_h \end{pmatrix}(V_r \; n_h) = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}. \tag{4.50}$$

Rearranging terms and using $Q(\xi_m^+) = g(Q(\xi_m^-))$, $\breve{\overline{X}}(\breve{\xi}_m^+, t) = g(\breve{\overline{X}}(\breve{\xi}_m^-, t))$ and (4.40), gives

$$\begin{aligned} \delta X(\xi_m^+) &\simeq -Q_\xi(\xi_m^+)\left(-\frac{\delta V(0)}{V_\xi(0^-)}\right) + \begin{bmatrix} 0 \\ n_{h_\xi}(0^-)\left(-\frac{\delta V(0)}{V_\xi(0^-)}\right) + \delta n_h(0^-) \end{bmatrix} \\ &= \begin{bmatrix} \frac{V_\xi(0^+)\delta V(0)}{V_\xi(0^-)} \\ \frac{n_{h_\xi}(0^+)\delta V(0)}{V_\xi(0^-)} - \frac{n_{h_\xi}(0^-)\delta V(0)}{V_\xi(0^-)} + \delta n_h(0^-) \end{bmatrix}. \end{aligned} \tag{4.51}$$

The above can be written as an expression for the perturbation after the unperturbed firing event $\xi = \xi_m^+ = m\Delta^+ = 0^+$ in terms of the perturbation before the unperturbed firing event $\xi = \xi_m^- = m\Delta^- = 0^-$:

$$\delta X(0^+) \simeq K_{\text{fire}}\delta X(0^-), \tag{4.52}$$

where

$$K_{\text{fire}} = \begin{bmatrix} 0 & 0 \\ \left(n_{h_\xi}(0^+) - n_{h_\xi}(0^-)\right)/V_\xi(0^-) & 1 \end{bmatrix}, \tag{4.53}$$

is called a "saltation matrix" and it describes how to propagate perturbations through the different domains. Note that $V_\xi(0^+) = 0$ because the voltage is clamped after the firing event.

The dynamics through the switching manifolds (from $\mu = 0$ to $\mu = +$) are continuous for the perturbed and unperturbed trajectories, i.e. $Q(\xi_s^+) = Q(\xi_s^-)$ and $\breve{\overline{X}}(\breve{\xi}_s^+, t) = \breve{\overline{X}}(\breve{\xi}_s^-, t)$, as can be seen in Figure 4.8. Therefore, g for these

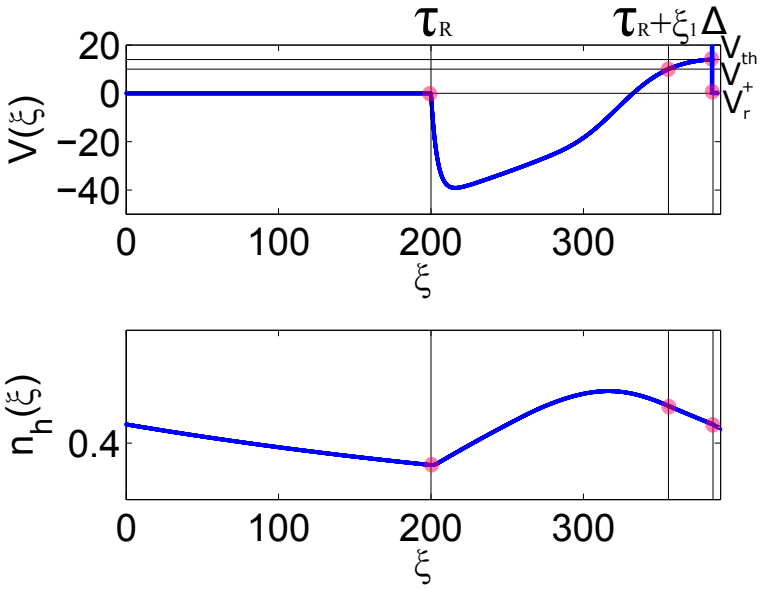


Figure 4.6. Evolution of the system (4.16) during one period. Top: V as a function of ξ. Bottom: n_h as a function of ξ. Red dots denote the points where the system changes dynamics, first from the refractory period at $\xi = \tau_R$ to the domain $\mu = 0$. Then it evolves until $V = V_+$ at $\xi = \tau_R + \xi_1$, when it switches to the domain $\mu = +$. There is a discontinuity in V when it reaches threshold at $V(\Delta) = V_{th}$ and is reset to $V = V_r$ for a duration of τ_R. Note that at this point n_h is continuous.

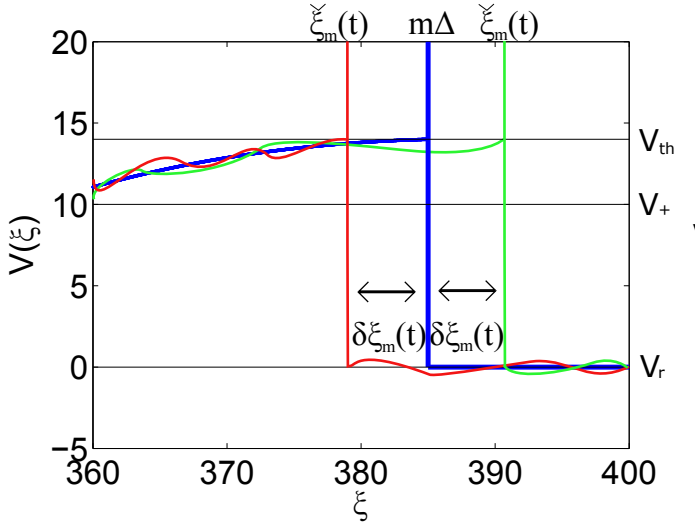


Figure 4.7. Unperturbed (blue line) and perturbed (red and green line) trajectories for V in (4.23) at the firing event. The unperturbed trajectory is the same as in Figure 4.6 (Top).

cases is defined by $g(\overline{X}(\xi,t)) = (\overline{V}(\xi,t), \overline{n}_h(\xi,t))^T$ (guaranteeing continuity), so that the gradient operator in (4.50) becomes the identity: $\nabla_Q^T g^T = I_2$. Similar calculations as for (4.48) and (4.49) yield

$$\delta X(\xi_s^+) \simeq -Q_\xi(\xi_s^+)\left(-\frac{\delta V(\xi_s)}{V_\xi(\xi_s^-)}\right) + Q_\xi(\xi_s^-)\left(-\frac{\delta V(\xi_s)}{V_\xi(\xi_s^-)}\right) + \delta X(\xi_s^-), \quad (4.54)$$

so that

$$\delta X(\xi_s^+) \simeq K_s \delta X(\xi_s^-), \quad (4.55)$$

where $K_s = I_2$.

The propagation of perturbations through the refractory state is harder to determine because this occurs over a finite time scale τ_R, unlike the firing events and switching that are instantaneous. As shown in Figure 4.9 the perturbation $\delta\xi$ at a firing event $\xi = \xi_m$ is propagated for a time τ_R before new dynamics are encountered with $\delta\xi$ given by (4.40). Setting $\xi_R = \xi_m + \tau_R$ as the time when the unperturbed trajectory changes dynamics, and $\check{\xi}_R = \check{\xi}_m + \tau_R$ as the time when the perturbed trajectory does, the continuity of a solution after releasing from refractoriness can be guaranteed by following the same reasoning as with the switching events with $\xi_m = 0$, yielding

$$\delta X(\xi_R^+) \simeq -Q_\xi(\xi_R^+)\left(-\frac{\delta V(0)}{V_\xi(0^-)}\right) + Q_\xi(\xi_R^-)\left(-\frac{\delta V(0)}{V_\xi(0^-)}\right) + \delta X(\xi_R^-). \quad (4.56)$$

Here $V_\xi(\xi_R^-) = 0$, since the system is in the refractory state, and $n_{h_\xi}(\xi_R^+) = n_{h_\xi}(\xi_R^-)$. Thus, substitution into (4.56) gives

$$\delta X(\xi_R^+) \simeq K_{\text{ref}} \delta X(0) + \delta X(\xi_R^-), \quad (4.57)$$

where

$$K_{\text{ref}}(\xi_s) = \begin{bmatrix} V_\xi(\xi_R^+)/V_\xi(0^-) & 0 \\ 0 & 0 \end{bmatrix}. \quad (4.58)$$

In conclusion, the perturbations can be mapped using saltation matrices as:

- $\delta X(0^+) = K_{\text{fire}} \delta X(0^-)$, through the firing event.
- $\delta X(\tau_R^+) = \delta X(\tau_R^-) + K_{\text{ref}} \delta X(0^-)$, through the refractory period.

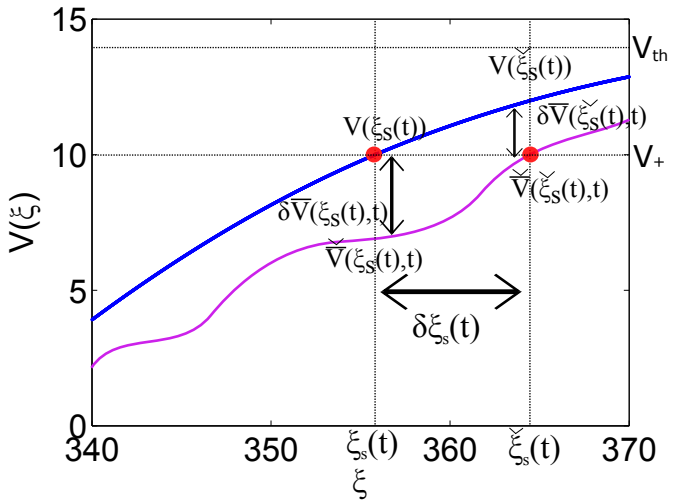


Figure 4.8. Unperturbed (blue line) and perturbed (purple line) trajectories for V in (4.23) at the switching event from the domain $\mu = 0$ to $\mu = +$ (red dots). Note the continuity of the trajectory through this switching. The unperturbed trajectory is the same as in Figure 4.6 (Top).

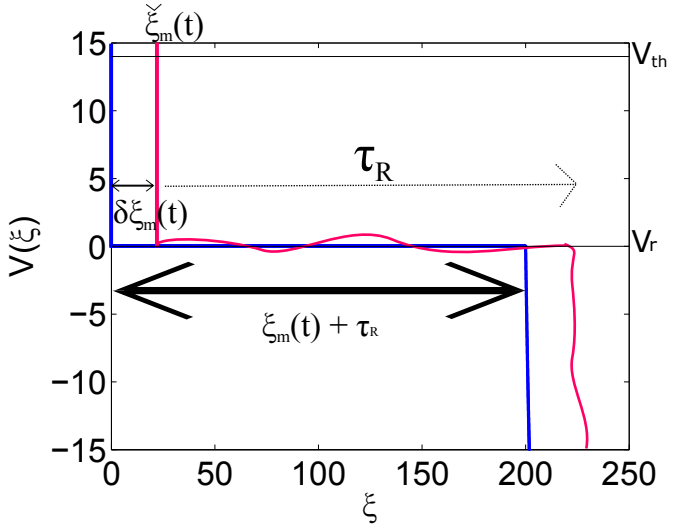


Figure 4.9. Unperturbed (blue line) and perturbed (pink line) trajectories for V in (4.23) showing the propagation of the perturbation $\delta\xi_m$ for $\xi = \tau_R$ when it is released from the refractory period and the dynamics change. Note the continuity of the trajectory through this switching. The unperturbed trajectory is the same as in Figure 4.6 (Top).

- $\delta X((\tau_R+\xi_1)^+) = K_s \delta X((\tau_R+\xi_1)^-)$, through the switch from the domain $\mu = 0$ to $\mu = +$. In general the same formula is used in the switches between $\mu = 0$ and $\mu = \pm$.

4.4.3 *Spectral Equation Determination*

The results of §4.4.2 provide a full description of the evolution of wave perturbations. Here the stability of the travelling waves is determined by exploiting this result. For convenience, the function $\mathcal{F}_\mu(\xi, \xi_0; \lambda)$ is introduced:

$$\mathcal{F}_\mu(\xi, \xi_0; \lambda) = \int_{\xi_0}^{\xi} \mathrm{d}\xi' G_\mu(\xi - \xi'; \lambda) J f(\xi'; \lambda), \qquad \mu \in \{0, +, -\}\,. \tag{4.59}$$

Combination of (4.59) and (4.45) allows a description of the perturbation after one period as $\delta X(\Delta) = \Gamma(\lambda, \Delta)\delta X(0)$, where

$$\begin{aligned} \Gamma(\lambda, \Delta) &= \mathcal{F}_+(\Delta, \tau_R + \xi_1; \lambda) \\ &+ G_+(\Delta - (\tau_R + \xi_1); \lambda) K_s \Big[\mathcal{F}_0(\tau_R + \xi_1, \tau_R; \lambda) \\ &+ G_0(\tau_R + \xi_1; \lambda) \left[G_R(\tau_R; \lambda) K_{\text{fire}} + K_{\text{ref}} \right] \Big]. \end{aligned} \tag{4.60}$$

The spectral condition $\varepsilon(\lambda, \Delta) = 0$ is obtained by enforcing the Δ-periodicity of the perturbations, here

$$\varepsilon(\lambda, \Delta) = \left| \Gamma(\lambda, \Delta) - I_2 \right|, \tag{4.61}$$

where $\varepsilon(\lambda, \Delta)$ is the Evans function for the periodic wave. As for the construction of the trajectory, a Fourier representation is used in (4.42) for numerical convenience. Thus $f(\xi; \lambda) = \sum_{p \in \mathbb{Z}} f_p(\lambda) \exp(-2\pi i p \xi / \Delta)$, where

$$f_p(\lambda) = -\frac{1}{V_\xi(0)} \frac{2\pi}{\Delta^2} i p \widetilde{\eta} \left(-i\lambda - \frac{2\pi p}{\Delta} \right) \widetilde{W} \left(\frac{2\pi p}{c\Delta} \right), \tag{4.62}$$

for $\mathrm{Re}\,(\lambda + \alpha) > 0$. Substituting in (4.59) and using properties of matrix exponentials gives:

$$\mathcal{F}_\mu(\xi, \xi_0; \lambda) = \sum_p f_p(\lambda) P_\mu \mathrm{diag} \left(S_\mu^+(\xi, \xi_0; \lambda), S_\mu^-(\xi, \xi_0; \lambda) \right) P_\mu^{-1} J, \tag{4.63}$$

where $\lambda_{\pm}$ is defined by (4.22) and

$$S_{\mu}^{\pm}(\xi,\xi_0;\lambda) = \frac{\mathrm{e}^{\left(\lambda_{\mu}^{\pm}-\lambda\right)(\xi-\xi_0)}\mathrm{e}^{-2\pi ip\xi_0/\Delta} - \mathrm{e}^{2\pi ip\xi/\Delta}}{\lambda_{\mu}^{\pm} - \lambda + 2\pi ip/\Delta}. \tag{4.64}$$

Figure 4.10 shows the eigenvalues of the system (green dots) when the wave is stable (Fig. 4.10a) and unstable (Fig. 4.10b). Consider the decomposition $\lambda = \nu + i\omega$ and $\varepsilon(\nu + i\omega, \lambda) = \mathcal{G}(\nu,\omega) + i\mathcal{H}(\nu,\omega)$, so that the eigenvalues are obtained by simultaneously solving the pair of equations $\mathcal{G}(\nu,\omega) = 0$ (red line) and $\mathcal{H}(\nu,\omega) = 0$ (blue line). Note that the system has a dynamic bifurcation for large Δ.

As expected, the travelling wave periods resulting from simulations are predicted to be stable. However, these periods are small compared to the maximum stable Δ determined by the stability analysis derived in this section. Due to the discrepancy between the maximum travelling wave periods resulting from simulations and from the stability analysis, a validation of the analytical results is presented in next section.

4.5 VALIDATION OF THE ANALYTICAL RESULTS

4.5.1 *Zero Eigenvalue*

Note that in Figure 4.10 there is always a zero eigenvalue. Hence it is important to stablish its persistence when Δ (or another parameter) varies. Differentiating (4.16) with respect to ξ gives

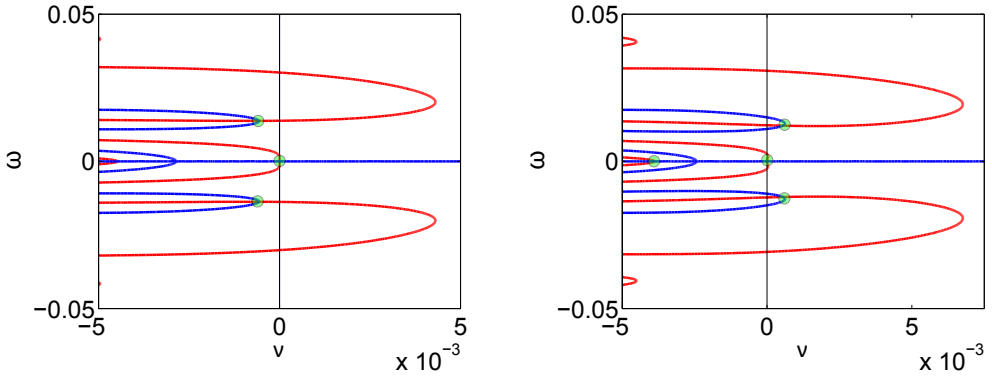
$$\frac{\mathrm{d}}{\mathrm{d}\xi}\frac{\mathrm{d}Q}{\mathrm{d}\xi} = A\frac{\mathrm{d}Q}{\mathrm{d}\xi} + \frac{\mathrm{d}\Psi(\xi)}{\mathrm{d}\xi}. \tag{4.65}$$

Combining (4.42) and (4.18) results in

$$\delta\psi(\xi;0) = \frac{\delta V(0)}{V_{\xi}(0)}\frac{\mathrm{d}\psi(\xi)}{\mathrm{d}\xi}. \tag{4.66}$$

Substituting the above equation in (4.43) with $\lambda = 0$ gives

$$\delta X(\xi) = \frac{\mathrm{d}Q}{\mathrm{d}\xi}. \tag{4.67}$$



(a) (b)

Figure 4.10. Zeros of the Evans function (4.61). These occur at the intersection (green dots) of $\mathcal{G}(\nu,\omega) = 0$ (red curve) and $\mathcal{H}(\nu,\omega) = 0$ (blue curve) where $\mathcal{G} = \mathrm{Re}\,(\varepsilon)$ and $\mathcal{H} = \mathrm{Im}\,(\varepsilon)$. (a) $\Delta = 460$, all the eigenvalues have $\nu \leq 0$, hence they are stable, whereas they are instable in (b) with $\Delta = 470$. Other parameters as in Figure 4.3.

Hence a perturbation tangential to the travelling wave orbit is neutrally stable, as expected from translation invariance of the system.

4.5.2 *Test against Simulations*

In this subsection, the accuracy of the stability analysis in §4.4 is confirmed by comparison with appropriate numerical simulations. The initial values for these simulations were taken from the theoretical travelling wave orbit at $t = 0$ and $x = -\xi/c$ in a domain sufficiently large to have periodicity two. It is assumed that neural activity was already spread through the network, so that some of the neurons had recently fired, and the time that these neurons had been refractory had therefore to be inferred. To avoid boundary effects the domain was taken to be periodic. Note that the spatial mesh size is crucial for numerical accuracy: small distance between neurons Δx, and therefore a large mesh, is required for the initial values of $X(x,t)$ at $t = 0$ so that the transformation from the theoretical travelling wave in ξ coordinates to (x,t) coordinates $x = -\xi/c$ is accurate (note that c values are small).

Figures 4.11a and 4.11c show that the simulated trajectories (red line) initiated with a travelling wave orbit corresponding to a neutrally stable period stay on (or close to) the theoretical orbit (blue line). On the other hand, Figures 4.11b and 4.11d show that the simulated trajectories diverge from the theoretical travelling wave orbit when initiated with a period predicted to be unstable.

4.6 PARAMETER CONTINUATION

This model provides a mechanism for the generation of periodic traveling waves, with a wavelength that depends on model parameters, and is given by the velocity of wave propagation c multiplied by the wave period Δ. Moreover, the model assumptions are biophysically reasonable. The work in [50] shows a difference in the I_h current time constant of stellate cells (that are believed to represent a large fraction of the grid cell population [46]) along the MEC dorsoventral axis, whereas [22,68] show that grid cells from more ventral portions of MEC II have larger grid field size compared to more dorsal portions. Figure 4.12a shows the dispersion curve for different values of τ_h. For small values of τ_h, the dispersion curve and the set of stable periods is small, whereas for large values of τ_h it increases. Note that the maximal stable period increases with the I_h current time constant τ_h, thus agreeing with experimental data from [50]. Figures 4.12b, 4.12c and 4.12d show the set of stable periods (blue) delimited by its maximum and minimum period (black line). Note that for the strength of I_h current, g_h, there is no significant change on the maximal period (Fig. 4.12d) unlike the time constant τ_h (Fig. 4.12b) and connectivity radius R (Fig. 4.12c). Although the scale of the travelling patterns can be controlled by varying τ_h, the maximal stable period for the travelling wave is more sensitive to changes in the connectivity radius than in the I_h time constant.

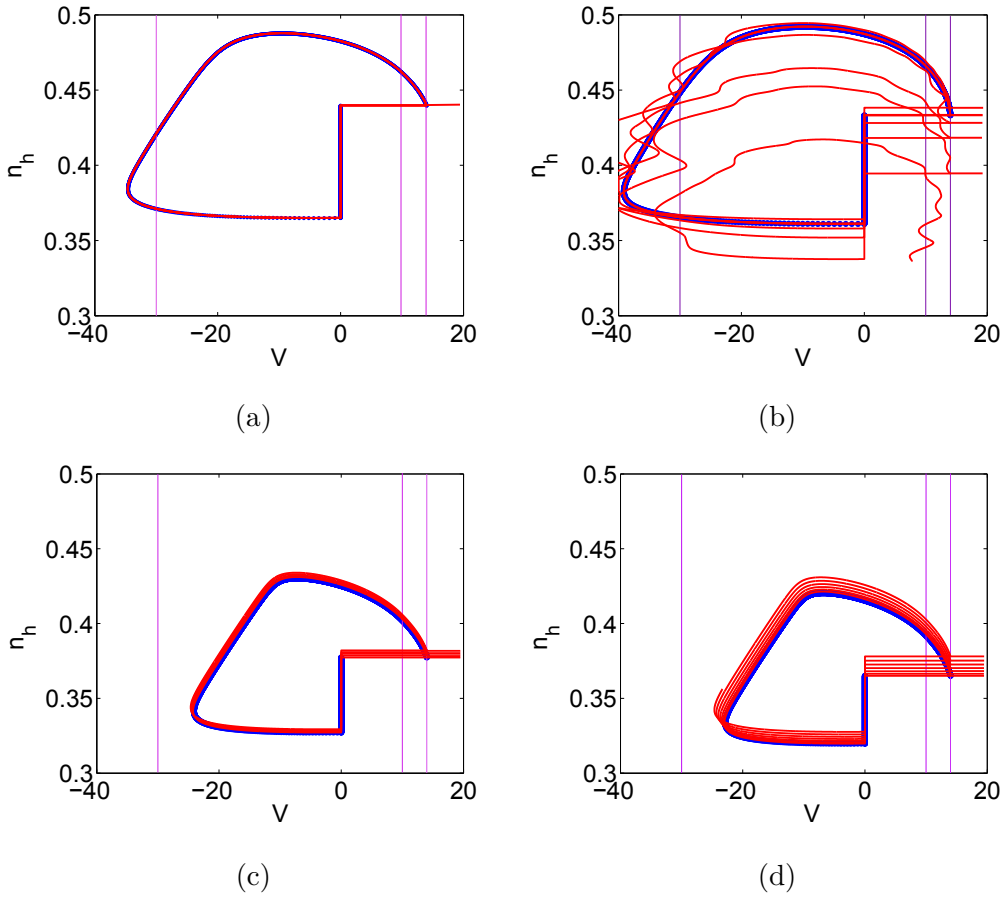


Figure 4.11. Orbits of the the travelling waves emerging from the IF network model: Theory (blue) (4.10) vs simulations (red) (4.8) for different periods. Here the simulations were initiated with the values corresponding to theory until $t = 3000$ ms. At the predicted stable periods ($\Delta = 385$ lower branch for (a) and $\Delta = 455$ for (c)), the simulated trajectories stayed close to the theory, whereas for the unstable periods ($\Delta = 385$ upper branch for (b) and $\Delta = 470$ for (d)) the simulated trajectories diverged from the theory (as expected). The unstable trajectories either evolve to a travelling wave with smaller period (d) or to random activity (b). Parameters as in Figure 4.3.

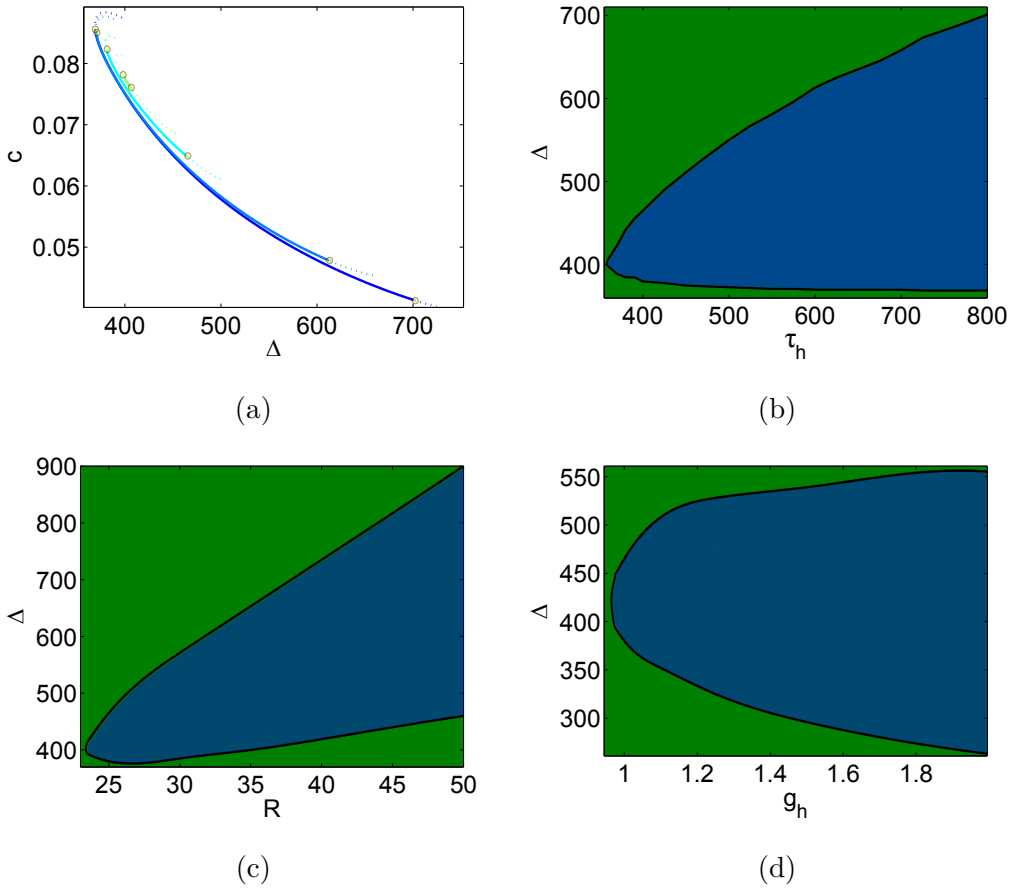


Figure 4.12. Continuation of travelling wave solutions as a function of various system parameters. (a) Dispersion curve $c = c(\Delta)$ for $\tau_h = 360, 400, 600, 800$ ms. The green line corresponds to $\tau_h = 360$ ms, cyan to $\tau_h = 400$ ms, light blue $\tau_h = 600$ ms and dark blue $\tau_h = 800$ ms. Solid lines represent periods where system is stable, while dashed lines represent where it is unstable. Yellow dots representing the values where the stability is lost. (b) Continuation of τ_h. This graph shows the maximum and minimum periods (black line) for which the system is stable for different values of τ_h. Thus, the blue zone correspond to the stable periods for a given τ_h (solid lines in (a)) whereas the green zone corresponds to the unstable periods. (c) and (d) same as (b) but for a given radius connectivity R and g_h respectively. Other parameters as in Figure 4.3.

4.7 CONCLUSIONS

Based on anatomical and electrophysiology constraints, a spiking network model that exhibits a wide range of long wavelength travelling wave solutions was proposed. To facilitate analysis a PWL function for the HCN channel activation function was used and a refractory period was imposed. After performing numerical simulations (details in Appendix F) travelling waves were observed whose orbit was described theoretically using techniques from the field of non-smooth dynamical systems. In order to determine the maximum allowed period, a wave stability analysis was developed. Importantly, using techniques from non-smooth dynamics, in §4.4.2 it was described how to handle the perturbations of the travelling wave orbit in time (by employing a co-moving framework) through the switching manifold discontinuities inherited from the HCN channels and the firing events. In the analysis presented in this chapter it was explained how these techniques can be extended, to evolve perturbations of the travelling wave through the refractory period, as well as be applied to switching and firing events. Using parameter continuation (§4.6), a relationship between the maximal period and the model parameters was established. Furthermore, it was demonstrated that the maximal period of the travelling wave increases with the I_h time constant τ_h which represents an alternative mechanism for modulating pattern wavelength independent of the connectivity radius. Note that unlike other spiking models of grid cell activity (and in general most of the LIF models for neural activity) that are explored predominantly with numerical simulations, the mathematical tractability of model proposed in this chapter allows a description of existence and stability, and therefore to predict the maximal stable period for a given parameter set.

4.7.1 *Further Work*

In the model proposed in this chapter, a refractory period was imposed to simplify the analytical description of the wave orbit. Hence, the model presented a single spike instead of the burst of spikes seen in the model proposed by Hasselmo [51]. The analysis presented in this chapter could be extended to the case where the refractory period is removed and the cell activity exhibits a burst of spikes, where all the spikes in the burst are contained within a period of the travelling wave.

The 1D spatial model simulations and analysis presented in this chapter has been published in Bonilla-Quintana *et al.* [69], along with additional simulations of the model in 2D. Thus a similar analysis to that presented in this chapter can be applied in order to predict the scale of emerging patterns. One of the major challenges to achieve this is to parametrise the shape of the 2D travelling wave.

The wave stability analysis described in §4.4 can also be applied to models with different types of neural connectivity or/and currents (with a corresponding PWL caricature of their activation function) due to its general form. Indeed, it would be interesting to extend the analysis to cover multiple currents with multiple gating variables[4], especially as K^+ and Na^+ currents have previously been taken into account in models based on experimental data that describe the resonance properties of grid cells [63].

4 Particularly, in the presented model the two currents refer to the synaptic and the I_h current.

5

A NAVIGATION MODEL USING FIRING FIELDS

5.1 INTRODUCTION

In Chapter 3 the relationship between the activity pattern formed by a continuous attractor neural network model and the resulting grid cell firing field was discussed. Nevertheless, the question of how these firing fields could be involved in navigation was not addressed. A firing field is obtained by recording the action potentials of a neuron while an animal is traversing an environment. The firing fields formed by grid cells and place cells are different. A grid cell fires when the rat is at different locations, resulting in a hexagonal firing field pattern, whereas a place cell just fires when the rat is at a certain spot. Figure 5.1 shows the firing fields from grid and place cells. In this figure the action potentials are represented as dots in the trajectory.

Since the discovery of grid cells, there has been a consensus that this type of cell is involved in the maintenance of estimates of current location by using self-motion signals to estimate travelled distances and directions [70]. Such computations are known as *path-integration*. Fiete *et al.* [71] noted that if grid cells were the hypothesised path-integrator, they must represent position with high resolution to minimise the accumulation of errors. They proved that a small number of grid cell firing fields with different spatial scales can uniquely specify the position on a large environment. As discussed in §3.1, grid cell firing fields increase their spatial scale from dorsal to ventral recording positions in the MEC II. Moreover, Stensola *et al.* [72] showed that such increments in

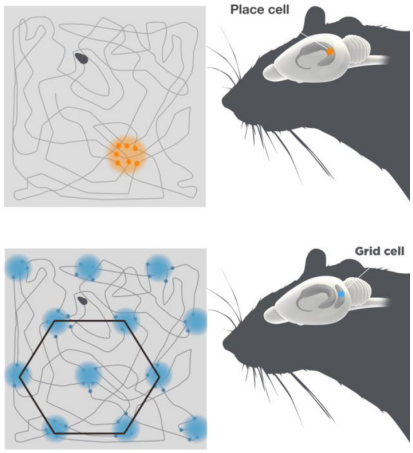


Figure 5.1. Firing Fields and anatomical locations of a place cells (Top) and grid cells (Bottom). Left: recorded action potentials of a place (orange dots) and a grid (blue dots) cell while a rat explores an environment. Here the gray line represents the rat's trajectory. Right: cell location in the rat brain. Place cells are found in the hippocampus and grid cells are found in the entorhinal cortex. Image by Mattias Karlen ©The Nobel Assembly at Karolinska Institutet, available at https://www.sciencenews.org/sites/default/files/images/fig1-2_advanced-medicineprize2014-2.png.

spatial scale are discrete. The total number of discrete modules formed by grid cells with common scale is hypothesised to be fewer than 10 [73], giving an experimental basis for the work in [71]. The importance of these discrete modules to navigation was recently studied by Wei *et al.* [74]. They proposed that the grid system implements a hierarchical code for space, where modules with small spatial scale provide precise location; whereas modules with large spatial scale resolve ambiguity. In this system, the modules are organised to minimise the number of neurons required to encode location with a given resolution across a range equal to the largest spatial period. Moreover, by using these assumptions, Wei *et al.* were able to explain the triangular lattice structure of two dimensional grid cell firing fields and predict a geometric progression of spatial scales.

Superficial layers of the MEC form a major spatial input to the hippocampus [75]. A natural question to ask is: how can the firing of grid cells in the MEC, when a rat is at multiple locations (forming a periodic grid in the envi-

ronment), be transformed into place cell firing that occurs when the rat is at a specific location (see Fig. 5.1)? In order to answer this question computational models that transform periodic grid cell firing fields into place fields have been proposed (see Cheng *et al.* [76] for a review). One example is the goal-directed navigation model[1] proposed by Erdem *et al.* [77] composed of a network of head direction cells, grid cells, place cells and prefrontal cortex cells. In this model place cells recieve synaptic input from a population of grid cells with different spatial scales. However, recent discoveries indicate that grid cells are not always crucial to place cell function [70]. Moreover, Poucet *et al.* [78] state that navigation based on grid cells alone (termed pure self-motion navigation) cannot remain accurate over indefinite distances or times, due to discrepancies between computed and true positions accumulating, unless the system resets to the true position using landmarks. They proposed a theory based on the idea that circuits controlling navigation receive information about the animal's location in its surroundings as well as self-motion. Interestingly, goal-directed navigation models were proposed before the discovery of grid cells, based only on place cell activity. Furthermore, these models are able to replicate experimental results. For example, Foster *et al.* [79] simulated experiments in which animals are able to find a hidden goal location from different starting positions. Moreover, their artificial animal reaches the goal location more rapidly in each trial as in experiments. In their work, Foster *et al.* employed an algorithm called temporal difference learning that is a method of reinforcement learning to solve the goal-directed task (see §5.2 for a short introduction).

Most goal-directed navigation models mimic maze experiments that were first performed by Tolman during the 1930s and 1940s in order to answer the question of how animals remember certain places. Roughly speaking, maze experiments consist of placing a rat within an enclosed maze environment, that additionally contains a reward (food) at a given (hidden) goal location. The rat must explore the environment in order to find the reward. In these experiments, the rat learns to take the quickest route towards the goal, thereby allowing

1 The objective of a goal-directed naviagation model is to find a (hidden) goal location.

investigation of navigation and learning processes in a controlled way. Tolman proposed that animals discover relationships between places and events as they explore the environment. Importantly, he suggested that exploration leads gradually to the formation of a *cognitive map* that enables the creation of detours and shortcuts in the presence of obstacles [80, 81]. It was not until 1971, with the discovery of place cells in the rat hippocampus by O'Keefe, that evidence was obtained that such a map structure of the spatial environment could indeed be formed in the brain [82].

In this chapter a navigation-based model using reinforcement learning is proposed. In order to outline the algorithms on which this model is based, first an introduction to reinforcement learning is given in §5.2. Then, the model is constructed and compared to previous models in the literature (§5.3). In §5.4 a new task is modelled in order to determine if the proposed model is capable of learning different goal locations, that change every day, after the first trial each day. Furthermore, in §5.4.2 the effects of adding landmarks to the environment are investigated, and in §5.4.3 the performance of the agent is studied under variation in the size and spacing of place cell firing fields.

5.2 REINFORCEMENT LEARNING

Here an introduction to reinforcement learning (RL), based on the seminal work of Sutton and Barto [83], is presented. These authors define RL as a computational approach to learning, based on the idea that humans (and animals) learn from interacting with the environment. In RL the learner (*agent*) is not told what to do, but instead discovers which actions give more reward by trying them. Note that this idea is similar to that proposed by Tolman for the formation of a cognitive map in the sense that the learning process occurs through interactions with the environment, and it is therefore natural to use reinforcement learning to simulate maze tasks. At the beginning of such tasks, the goal location is unknown for the rat, so it explores the environment until it reaches

the goal location where it receives a reward. On the subsequent trials[2], the rat will have some information about the goal location that influences its decision of where to go (action) in order to reach the goal location more efficiently.

In order to state a general maze task in a reinforcement learning framework, first some key concepts must be defined. Here, the rat is referred to as the agent: it takes the decisions in view of its current *state* (here defined by its position in the maze). A *policy* defines the learning agent's way of behaving at a given time. It can be seen as a map linking previously explored locations to actions to be taken when in these locations. On the other hand, a *reward function* maps each state (or state-action pair) of the environment to a single number; a reward. The reward function is given by the environment in the sense that the agent cannot change it. Instead the agent changes its policy based on the reward at each time step. At each time step, the agent estimates a *state-value function* that gives a notion of "how good" it is for an agent to be in a given state. This value function is the total amount of reward an agent can expect to accumulate over the future, starting from its current state. The agent also estimates "how good" it is to choose an action in a given state with an action-value function for a given policy. Note that RL is based on three processes: *sensation* (information about the environment), *action*, and *goal*. Therefore, it can be defined as learning how to map situations to actions, so as to maximise the reward signal it receives in the long run.

In more formal terms, the agent and the environment interact at every time point t, the agent receives some representation of the environment's state $s_t \in \mathcal{S}$, where $\mathcal{S}$ is the set of possible states, and based on that, it selects an action $a_t \in \mathcal{A}(s_t)$, where $\mathcal{A}(s_t)$ is the set of available actions at state s at time t. As a consequence of this action, the agent transitions to a new state s_{t+1} and receives a reward $r_{t+1} \in \mathbb{R}$ on the next time step. The policy $\pi_t(s, a)$ is defined as the probability of taking the action a at time t if the agent is on state s. The state-value function for state s, $V^\pi(s)$, is the expected accumulative

2 A trial begins when the rat is put on the environment and finishes when the rat finds the goal location and obtains a reward.

reward when starting in s and following π thereafter. The action-value function, $Q^{\pi}(s, a)$, is the expected accumulative reward starting from s, taking action a, and thereafter following policy π. In RL, the actions taken by the agent may not only affect the immediate reward but also the subsequent situation, that could lead to subsequent rewards. This gives rise to the notion of a *delayed reward*. For this reason such models are often termed temporal difference learning.

An example of an artificial maze experiment on a square lattice, where every square represents a position, is presented in Figure 5.2 in order to illustrate the above. A trial starts when the agent is put on a square and ends when it reaches the goal location and receives a reward. In this maze example the reward function is set to the value 1 on the goal location and -1 everywhere else, so the agent is encouraged to reach the goal location as fast as possible from any location. Since the goal location is hidden, the agent can not take a direct path towards it. Instead, it must explore the environment until it reaches the goal location. The artificial rat makes a decision of where to move at each time step based on previously obtained rewards, or it can make a random movement with the aim to explore other areas. One of the challenges of this RL approach is the trade-off between exploration and exploitation: actions that were already taken and led to a reward must be repeated, but also unexplored actions must be taken in order to discover other rewards. As indicated in Figure 5.2, the agent is constrained to a discrete lattice and it can choose to move only north, south, west or east. Dark blue arrows indicate the "chosen" path; other example paths are shown for comparison. From the starting position the agent could have chosen a better route by initially travelling east instead of north (light blue arrows). The advantage of reinforcement learning is that the agent can explore options leading to better routes, such as the one shown by purple arrows. Here the agent, instead of following a known route (dark blue arrows), chooses to turn left after three time steps.

In Figure 5.2, after reaching the goal at $s_N = \text{goal}$, the agent's policy changes on the previous location s_{N-1} to go south because it is the action that yields the most reward (on the dark blue route). The estimated value function decreases

with the distance of the goal location: the rat first explores, reaches the goal (dark blue arrows route), and back-propagates the information about the goal by assigning some discounted value to the proximal states. Now, suppose that in the next trial the rat follows the same path (purple route). Although it previously reached the goal location, the value assigned to going north on the third time step is very low (it took many steps to reach the goal), so the rat chooses to take another action in order to gain more reward. The goal location is then reached in three steps instead of seven, increasing the value of that location. The policy is changed accordingly. Note that the calculation of the value functions $V^\pi(s)$ and $Q^\pi(s,a)$ at any time requires knowledge of the reward function of the future states which (in this maze example) is unknown by the agent. Therefore an estimation is used instead. Finding the estimation $v(s)$ of $V^\pi(s)$ is the so-called *prediction problem* in RL, whereas improving the policy defines the *control problem*. The estimation of the value function is updated at every time step.

5.2.1 *Temporal Difference Learning*

Temporal difference (TD) methods learn directly from raw experience by updating the estimate state-value $v(s)$, based on a difference between two estimates at two different times, without waiting for a final outcome (i.e. they *bootstrap*). For now, the estimated state-value function is given by a number, but it also can be a parametrised function, which will be shown in §5.2.2. Imagine that the agent is at location s at time t, s_t, where it takes an action that leads it to state s_{t+1}. In this new state the agent observes a reward r_{t+1}. Then the value at the earlier location $v(s_t)$ is adjusted to be closer to the value of the later state $v(s_{t+1})$ given the reward:

$$v(s_t) \leftarrow v(s_t) + \alpha \left[r_{t+1} + \gamma v(s_{t+1}) - v(s_t) \right], \tag{5.1}$$

where $0 \leq \gamma \leq 1$ is the discount rate and $\alpha > 0$ is a small, so-called, step-size parameter, which influences the rate of learning. Note that the value update

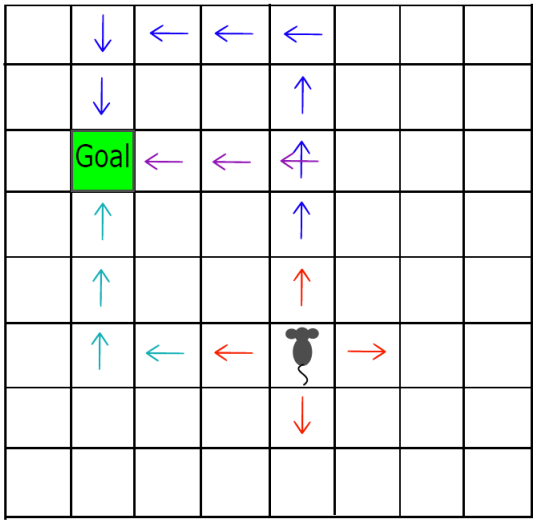


Figure 5.2. A general maze task. The rat (agent) is constrained to a discrete version of a maze (environment). On each square (location) the rat perceives its environment and makes a decision on which action to take. The allowed actions are going north, south, east or west (red arrows). The aim of the task is to learn which actions to take in order to minimise the number of movements required to reach the goal location from any location. Dark blue, purple and light blue arrows indicate different example paths from the agent's current position to the goal.

is based on an existing value estimate of the previous locations. For example, let $s_1 = \text{start} \rightarrow s_2 = B \rightarrow s_3 = C \rightarrow \ldots \rightarrow s_N = D \rightarrow s_{N+1} = \text{goal}$ denote the sequence of locations that the agent followed on the first trial (in Fig. 5.2 each location is represented as a square in the lattice). At the start of this trial, the estimated value function is similar for all the states because it is assumed that the environment is unknown by the agent. Consequently, the estimated value function does not change until the agent receives a reward. For simplicity, assume that only the goal location has a (positive) reward. After reaching the goal location the agent updates the estimated value of the previous state $v(D)$ by some fraction of the reward r_{N+1} obtained at the goal location using (5.1). Thus, the estimated value of state D is larger than the rest of the states. The policy $\pi_N(D, a)$ is changed accordingly by increasing the probability of taking the action a that led to the goal location. On the following trial, the rat follows the state sequence $s_1 = \text{start} \rightarrow s_2 = B' \rightarrow \ldots \rightarrow s_M = C' \rightarrow s_{M+1} = D$. Because the estimated value function at state D is larger than the one at state C', $v(C')$ increases and is larger than for the rest of the states but $v(D) > v(C')$, in this way, the information of the goal location is back-propagated. Also, it is more probable that the agent takes the action that leads to the goal location from state D due to the change of policy on the first trial.

5.2.1.1 *Actor-Critic Methods*

The TD learning method gives a solution for the prediction problem in RL. Nevertheless, the solution for the control problem has not yet been addressed. There are many TD methods that solve such problem based on the idea that the policy and the estimated value function updates should interact in such a way that both of them move towards the optimal value. For example, the *actor-critic method* has a separate memory structure to represent explicitly the policy, independent of the state-value function [83]. The policy structure is known as the *actor* and the estimated state-value function is known as the *critic*. Hence, the critic criticises the actions selected by the actor. The learning is *on-policy*: the critic learns about and criticises the policy followed by the actor at each

time step. The critic employs a *TD error* to evaluate if the action taken, a_t, by the actor at state s_t is better or worse than expected. The TD error is given by:

$$\delta_t = r_{t+1} + \gamma v(s_{t+1}) - v(s_t). \tag{5.2}$$

Note that δ_t is the change in the estimated state-value function in (5.1). Therefore, if the error is positive then the tendency to select a_t should be strengthened, whereas if the TD error is negative, it suggests that the tendency should be weakened.

5.2.1.2 *Eligibility Traces (TD(λ) Method)*

So far, the state-value function update (5.1) considers only the values of the previous and the current states. However, a temporary record of the occurrence of visiting a state or of taking an action can also be implemented. Such a record is called an *eligibility trace*. For example, let the event be visiting a location in the maze task example in Figure 5.2, then each time that the agent visits that particular location, the eligibility trace will increase, otherwise it decreases. This kind of eligibility trace is called an *accumulating trace* because it accumulates the times that a state has been visited by the agent, then fades away gradually when the state is not visited, as shown in Figure 5.3.

The TD(λ) algorithm implements an eligibility trace. Let $e_t(s) \in \mathbb{R}^+$ denote the eligibility trace for state s at time t. On each step, the eligibility trace for

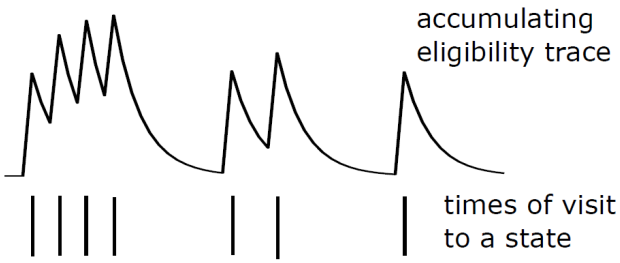


Figure 5.3. Accumulating eligibility trace for a state over time. On each time step, the eligibility trace decays by $\gamma\lambda$ if the state is not visited, if the state is visited it is incremented by 1.

the current state is incremented by 1, and the eligibility traces for all other states decay by $\gamma\lambda$:

$$e_t(s) = \begin{cases} \gamma\lambda e_{t-1}(s), & \text{if } s \neq s_t \\ \gamma\lambda e_{t-1}(s) + 1, & \text{if } s = s_t \end{cases}, \qquad \forall s \in \mathcal{S}, \tag{5.3}$$

where γ is the discount rate in (5.2). Henceforth, λ, $0 \leq \lambda \leq 1$ is the *trace-decay* parameter. The TD error signal is given by (5.2) and the state-value update to all recently visited states (i.e., $e_t(s) \neq 0$) is proportional to the TD error

$$v_t(s) \leftarrow v_t(s) + \alpha\delta_t e_t(s), \qquad \forall s \in \mathcal{S}, \tag{5.4}$$

where "recently" is defined by $\gamma\lambda$.

5.2.2 *Generalisation and Value Function Approximation*

In the general maze task (Fig. 5.2) most locations will have never been experienced, and the only way to learn anything at all is to generalise from previously explored locations to the locations that have never been seen. Thus, a combination of reinforcement learning methods with a generalisation method is needed. In particular, here the sparse state-value function $V^\pi(s)$ data obtained from visited locations is employed to approximate this function over the entire domain. The approximate value function at time t, V_t, will be represented as a parameterised functional form with parameter w_t, i.e. $V_t = V_t(s; w_t)$. In this way V_t depends on w_t, which varies from time step to time step. The mean-square error between V_t and the true value function V^π over some distribution (weighting the errors of different states), P, is given by

$$MSE(w_t) = \sum_{s \in \mathcal{S}} P(s) \left[V^\pi(s) - V_t(s; w_t)\right]^2. \tag{5.5}$$

This can be minimised using w_t if it is assumed that V_t is a smooth differentiable function of w_t for all $s \in \mathcal{S}$, where w_t is a column vector, and that on each time step there is full knowledge of the corresponding $V^\pi(s_t)$. A good strategy is to try to minimise the error between the state-value function and the approximate

value function for the corresponding state at time t. *Gradient-descent methods* do this by adjusting the parameter vector, after the agent arrives at state s at time t and computes $V^\pi(s_t)$, by a small amount in the direction that would most reduce the error on that step:

$$\begin{aligned} w_{t+1} &= w_t - \frac{1}{2}\alpha\nabla_{w_t}\left[V^\pi(s_t) - V_t(s_t; w_t)\right]^2 \\ &= w_t + \alpha\left[V^\pi(s_t) - V_t(s_t; w_t)\right]\nabla_{w_t}V_t(s_t; w_t), \end{aligned} \tag{5.6}$$

where α is a positive step-size parameter[3], and $\nabla_{w_t}V_t(s_t; w_t)$ denotes the vector of partial derivatives of $V_t(s_t; w_t)$ with respect to w_t. Note that w_t is proportional to the negative gradient of the observed $V^\pi(s_t)$ squared error. This is the direction in which the error falls most rapidly.

As discussed in §5.2, the agent needs to use an estimation of the state-value function $V^\pi(s_t)$, $v(s_t)$. Thus, the general method for state-value prediction is given by substitution into (5.6) to yield:

$$w_{t+1} = w_t + \alpha\left[v(s_t) - V_t(s_t; w_t)\right]\nabla_{w_t}V_t(s_t; w_t). \tag{5.7}$$

Note that in (5.7) an estimation, v, of the state-value function, V^π, for the state s at time t is combined with a generalisation of $V^\pi(s_t)$ to other unexplored states, $V_t(s_t; w_t)$. A good estimation of $V^\pi(s_t)$ is achieved using the TD(λ) method when the TD error is zero, which using (5.2) gives $v(s_t) = r_{t+1} + \gamma v(s_{t+1})$. Now, assuming that the estimated value at state s and time t is given by a parametrised function instead of just a number, i.e. $v(s_t) = V_t(s_t; w_t)$, the learning process occurs through parameter updates at each time step:

$$w_{t+1} = w_t + \alpha\delta_t e_t, \tag{5.8}$$

where δ_t is the TD error given by

$$\delta_t = r_{t+1} + \gamma V_t(s_{t+1}; w_t) - V_t(s_t; w_t). \tag{5.9}$$

Here e_t is a column vector of eligibility traces, one for each component of w_t, updated by

$$e_t = \gamma\lambda e_{t-1} + \nabla_{w_t}V_t(s_t; w_t), \tag{5.10}$$

3 Note that α is also in (5.1) and in both equations denotes the amount of change that is assigned to the updated value at each time step.

with $e_0 = 0$. Note that the increments on the eligibility trace differs from that in (5.3), since the increments are proportional to the gradient of V_t and thus generalise to other states.

5.2.3 *Coarse Coding*

Until now, locations have been represented as a square in the discrete lattice environment (Fig. 5.2). However, locations in a 2D environment can have a continuous representation as a vector with two real components using *coarse coding*. In coarse coding, a state is represented with features that overlap. For example, in the artificial maze the environment can be represented by overlapping circles instead of a lattice (see Fig. 5.4). The overlapping circles are the features and they represent the agent's location by determining whether a feature is present in the current state or not. In Figure 5.4 a feature is present if the agent is at a location inside a particular circle or not. For example, when the agent is at state X there are three features present, and if it moves to state Y on the next time step, only one of these features remains present. Note that the representation of the states with features allows the agent to move in a continuous state space. Instead of moving to a different square in a lattice, it moves to different states where some of the features are shared between them. This allows a generalisation between states.

To formalise the above, let $\phi_{s,i}$ represent the value of the ith feature present at state s, then the approximate state-value function can be given by

$$V_t(s; w_t) = \sum_{i=1}^{n} w_{t,i} g(\phi_{s,i}), \tag{5.11}$$

where ϕ_s is a column vector with the same number of components as w_t and g is a function. Note that if $g(\phi_s) = \phi_s$ then $V_t(s)$ is linear in the parameters. There are different types of feature depending on the value assigned to represent the features. For example, a *binary feature* (see Fig. 5.4) assigns a value of 1 if the corresponding feature is present at the current state and 0 otherwise.

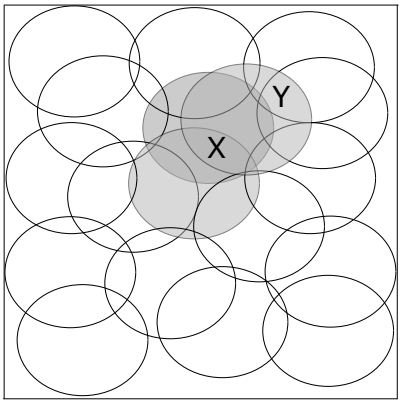


Figure 5.4. Coarse coding. There are three features present at state (location) X represented by three overlapping gray circles. Note that state X and state Y have one feature in common, so there will be slight generalisation between them.

5.2.3.1 *Radial Basis Functions*

Other type of features can be continuous-valued reflecting various degrees to which the feature is present. This type of features are called *radial basis function* (RBF) features, and typically these features have a Gaussian response that depends only on the distance between the state and the feature centre c_i, which is relative to the feature's width σ_i

$$\phi_{s,i} = \exp\left(-\frac{\|s - c_i\|^2}{2\sigma_i^2}\right). \tag{5.12}$$

Importantly, some theoretical models of place field formation also use Gaussian functions to represent place cell firing fields [84]. Hence, it is natural to use RBF features in the maze.

5.2.4 *Reinforcement Learning in Continuous Time and Space*

Doya [85] considered a more realistic approach to learning, in which TD methods are formulated for continuous time nonlinear dynamical systems instead of discrete time step iterations. This approach is more suitable for the maze task

because it allows the agent to take continuous actions in response to real valued sensory input. Doya considered a continuous time deterministic system,

$$\dot{s}(t) = f(s(t), a(t)), \tag{5.13}$$

where $s \in \mathcal{S} \subset \mathbb{R}^n$ is the state and $a \in \mathcal{A} \subset \mathbb{R}^m$ is the action. Here $\dot{}$ denotes the derivative with respect to time. The immediate reward for the state and the action is given by

$$r(t) = r(s(t), a(t)). \tag{5.14}$$

The goal of this formulation is to find a policy,

$$a(t) = \pi(s(t)), \tag{5.15}$$

that maximizes the cumulative future reward,

$$V^{\pi}(s(t)) = \int_t^{\infty} \mathrm{e}^{-\frac{t'-t}{\tau}} r(s(t'), a(t'))\mathrm{d}t' \tag{5.16}$$

for any initial state $s(t)$. Here τ is the time constant for discounting future rewards. The value function V^* for the optimal policy π^* is defined as

$$V^*(s(t)) = \max_{a[t,\infty)} \left[\int_t^{\infty} \mathrm{e}^{-\frac{t'-t}{\tau}} r(s(t'), a(t'))\mathrm{d}t' \right]. \tag{5.17}$$

The above integral can be evaluated using the principle of optimality. This principle was described by Bellman [86] as: "an optimal policy has the property that whatever the initial state and initial decision are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision". According to this principle the integral in (5.17) can be divided into two parts $[t, t + \Delta t]$, corresponding to the initial decision, and $[t + \Delta t, \infty)$. After solving the resulting integrals and taking $\Delta t \to 0$, the condition for the optimal value function at time t is found:

$$\frac{1}{\tau} V^*(s(t)) = \max_{a(t) \in \mathcal{A}} \left[r(s(t), a(t)) + \frac{\partial V^*(s(t))}{\partial s(t)} f(s(t), a(t)) \right], \tag{5.18}$$

which is the Hamilton-Jacobi-Bellman (HJB) equation for infinite-horizon, discounted reward problems and is used to solve optimization problems in continuous time systems. The optimal policy is given by

$$a(t) = \pi^*(s(t)) = \mathrm{argmax}_{a \in \mathcal{A}} \left[r(s(t), a) + \frac{\partial V^*(s(t))}{\partial s(t)} f(s(t), a) \right]. \tag{5.19}$$

The change of the value function over time under the policy π is given by differentiating (5.16) with respect to time, which yields $\dot{V}^{\pi}(s(t)) = V^{\pi}(s(t))/\tau - r(t)$. Because states and actions are continuous, the state-value function cannot be computed exactly over the entire domain (since all locations cannot be visited; see §5.2.2). Therefore, a value function approximation is needed. Let $V(s(t); w)$ denote the parametrised approximation of the state-value function, then the continuous version of the TD error is given by

$$\delta(t) = r(t) - \frac{1}{\tau}V(s(t)) + \dot{V}(s(t)). \tag{5.20}$$

As in the discrete case, the objective is to bring the TD error to zero. This can be achieved by implementing a gradient descent algorithm to minimise $|\delta|^2$ where

$$\dot{w_i} = \eta\delta(t)\left[\frac{1}{\tau}\frac{\partial V(s;w)}{\partial w_i} - \frac{\partial}{\partial w_i}\left(\frac{\partial V(s;w)}{\partial s}\right)\dot{s}(t)\right]. \tag{5.21}$$

Here η is a learning rate.

In order to derive a continuous version of the eligibility traces, Doya considered an impulse of reward at time $t = t_0$ in (5.16), to give

$$V^{\pi}(t) = \begin{cases} \mathrm{e}^{-\frac{t_0 - t}{\tau}}, & t \leq t_0 \\ 0, & t > t_0 \end{cases}. \tag{5.22}$$

Because the value function is linear with respect to the reward, the correction of the value function for an instantaneous TD error $\delta(t_0)$ (Fig. 5.5) is

$$V_c(t) = \begin{cases} \delta(t_0)\mathrm{e}^{-\frac{t_0 - t}{\tau}}, & t \leq t_0 \\ 0, & t > t_0 \end{cases}. \tag{5.23}$$

Therefore, he proposed the following update of w_i given $\delta(t_0)$

$$\dot{w_i} = \eta\delta(t_0)\int_{-\infty}^{t_0} \mathrm{e}^{-\frac{t_0 - t}{\tau}}\frac{\partial V(s(t);w)}{\partial w_i}\mathrm{d}t \equiv \eta\delta(t_0)e_i(t). \tag{5.24}$$

Note that the above equation reflects an eligibility trace in the sense that the exponential term fades as $t \to -\infty$. Hence the influence of the reward pulse is higher for previously visited states. In this way information about the reward

location is back-propagated as in the discrete case. The continuous version of TD(λ) is given by

$$\dot{w}_i = \eta\delta(t)e_i(t), \tag{5.25}$$

$$\dot{e}_i(t) = \frac{1}{\kappa}e_i(t) + \frac{\partial V(s(t);w)}{\partial w_i}, \tag{5.26}$$

where $0 < \kappa \leq \tau$ is the time constant for the eligibility trace.

Doya also developed a continuous version of the actor-critic algorithm where the TD error is used to improve the policy. He considered the following policy implemented by the actor:

$$a(t) = g\left(A(s(t);w^A) + \sigma_n n(t)\right), \tag{5.27}$$

where $A(s(t);w^A) \in \mathbb{R}^m$ is a function approximation with parameter vector w^A, $n(t) \in \mathbb{R}^m$ is noise, and g is a monotonically increasing output function. The size of the perturbation σ_n decreases as the performance improves; and the parameters are updated as:

$$\dot{w}_i^A = \eta_A\delta(t)n(t)\frac{\partial A(s(t);w^A)}{\partial w_i^A}. \tag{5.28}$$

The inclusion of the noise term allows the agent to explore different actions for every state, as in Gullapalli's stochastic real valued unit algorithm [87]. The actions become more deterministic (i.e., the noise decreases) as the agent learns about the best route towards the goal location; and therefore allows a good balance between exploitation and exploration.

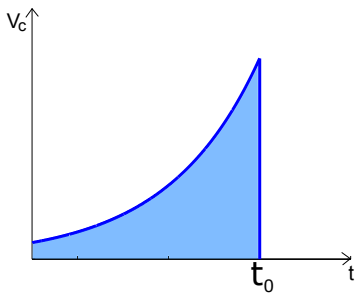


Figure 5.5. Update for the value function estimate $V_c(t)$ for an instantaneous TD error $\delta(t)$ at $t = t_0$. Here the instantaneous TD error is corrected by an exponentially weighted increase in $V(t)$ for $t < t_0$.

5.3 MAZE TASK USING FIRING RATE NEURAL NETWORKS

Frémaux *et al.* [88] proposed a model of a maze task with reinforcement learning using a continuous time actor-critic framework with spiking neurons. Their model is based on experimental evidence that dopamine, a neurotransmitter associated with pleasure, is released in the brain when a reward event occurs. Dopamine has also been shown to modulate spike-timing-dependent plasticity, and hence their choice of a spiking network. However, there is experimental evidence that dopamine is related to the state of desire (i.e., wanting a reward) rather than pleasure (i.e., liking a reward) [89].

Figure 5.6 shows a description of the model architecture proposed by Frémaux *et al.* in [88]. The actor and the critic are represented by two neural populations receiving information about the environment via place cell activity. The critic network computes the TD error that would strengthen or reduce the connections between place cells and the actor/critic neurons. The maze in Frémaux *et al.* has a square shape, with a U-shaped obstacle around the goal location, which is in the maze centre. The objective of the task is for the agent to learn the shortest path towards the goal, for any of the four starting locations.

In this section, a new model of this maze task is proposed using firing rate descriptions for the critic and the actor. Implementing firing rate instead of spikes decreases the computational cost of the simulations and facilitates mathematical analysis. In this model a trial starts when the agent (an artificial rat) is put at an initial state (a location in the maze) $\mathbf{x} \in \mathbb{R}^2$, $\mathbf{x} = (x, y)$. After the trial starts, the jth place cell's instantaneous firing rate is updated to

$$\rho_j^P(\mathbf{x}(t)) = \rho_P \exp\left(-\frac{\|\mathbf{x}(t) - \mathbf{x}_j\|^2}{\sigma_P^2}\right), \tag{5.29}$$

where ρ_P is constant and $\mathbf{x}_j$ is the location of the firing field centre encoded by the jth place cell, $j = 1, 2, \ldots, N_P$. The firing field centres are arranged on a uniform grid. Note that RBF features are used to represent the firing rates of place cells. The artificial rat senses its position in the environment only through place cell information. The presynaptic activity in place cells

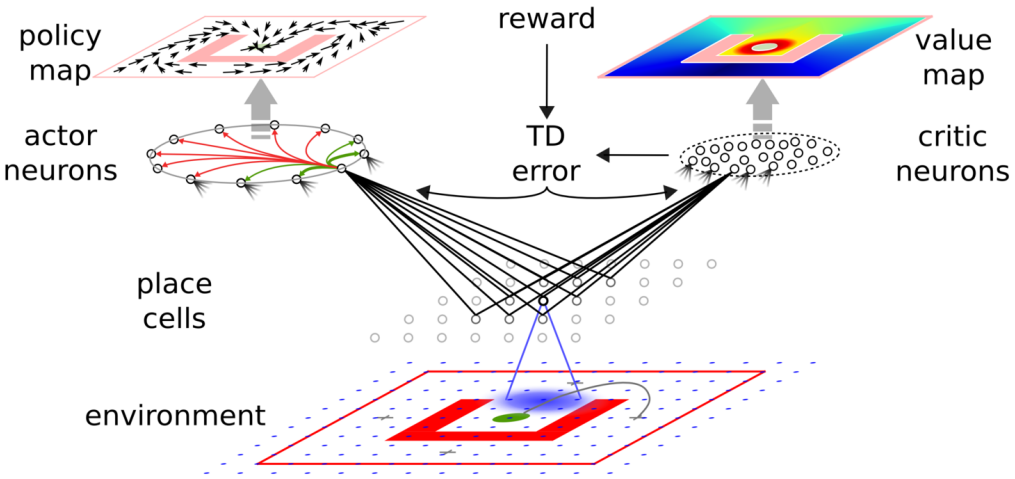


Figure 5.6. Maze task simulation by Frémaux [88]. Bottom: maze environment. Blue dots represent the centres of the place cell fields. Blue shadow shows how the activity of a place cell decays as the distance from the centre of its firing field increases (5.29). Grey line: evolution of the agent from the starting position (blue crosses) until it reaches the goal location (green disk). During the trial the agent avoids the obstacles (red U-shaped obstacle around the goal location and red lines delimiting the maze). Centre right: critic neurons that compute the TD error. Top right: expected future reward (value map) encoded by the critic. Centre left: actor neurons connected in such a way that a neuron excites nearby neurons while inhibiting all other neurons. Top left: the most probable actions that the agent can take (policy map). Figure taken from [88].

generates activity in the postsynaptic neurons of the critic and the actor. The agent receives a reward R when it reaches the goal location and is punished (via negative reward) when it runs into a wall. Thus, the reward R only occurs at precise times. However, to motivate the agent to achieve the goal as quickly as possible there must be temporal decay of the reward. This is modelled by a decaying reward rate function $r(t)$ that instantaneously updates when a reward R happens.

Importantly, because the rate model proposed here is deterministic it allows for the implementation of eligibility traces for the learning rule, as proposed by Doya [85], instead of having to simulate the so-called "neutral state" proposed by Fremaux *et al.* [88], in which learning continues for three seconds after reaching the goal location without place cell input. This "neutral state" resembles an eligibility trace and was used by Fremaux *et al.* to overcome problems associated with the inherent noise of spiking neurons.

5.3.1 *The Critic*

As stated previously, the value function estimation is performed by the critic. Here the critic is represented by a population of N_C independent neurons whose activity, $u_i^C(t)$, evolves according to

$$Q^C u_i^C(t) = \varepsilon_0 \sum_{j=1}^{N_P} w_{ij}^{PC}(t)\rho_j^P(t), \qquad Q^C = \left(1 + \tau_m \frac{\mathrm{d}}{\mathrm{d}t}\right)\left(1 + \tau_s \frac{\mathrm{d}}{\mathrm{d}t}\right), \quad (5.30)$$

where ε_0 is a scaling constant and $i = 1, \ldots, N_C$. Here a matrix $w^{PC} \in \mathbb{R}^{N_C \times N_P}$ is used for the parameters instead of a vector since the critic is represented as a population of neurons. Hence w_{ij}^{PC} represents the connection between the jth place cell and the ith critic neuron. For the firing rate function ρ^C, an exponential function is chosen

$$\rho_i^C(u_i^C) = \mu \exp(\beta(u_i^C - \Theta)). \quad (5.31)$$

The value function is given by

$$V(\mathbf{x}(t)) = \frac{v}{N_C} \sum_{i=1}^{N_C} \rho_i^C(t) + V_0, \quad (5.32)$$

where v and V_0 are constants. Here the value function is a smooth function of the place cell activity $V(\mathbf{x}(t); w^{PC}) = V(\rho^C(u^C(\rho^P(\mathbf{x}(t)); w^{PC})))$, and therefore a nonlinear function of the RBF features. The TD error and the parameter updates using the proposed rate networks and elegibility traces are derived in Appendix G. The parameter updates can be interpreted as a form of Hebbian-TD[4] error rule in the sense that presynaptic activity of place cells induces a change in the postsynaptic activity of critic neurons through a reward modulated signal.

5.3.1.1 *Linear Track Simulation*

In order to test if the proposed rate model can achieve similar results to the spiking neuron model in [88], a linear track simulation is performed. In this set up, the maze in Figure 5.6 is replaced with a rectangular track; with the goal placed at some distance ahead of the agent's starting position. The agent's choice is therefore clamped to the action of going straight towards the goal (and so actor neurons are not simulated). This simplified scenario allows a study of the critic learning in isolation. At the beginning of each trial, the artificial rat is placed in the maze at the same starting position from which it will run straight to the goal location. The trial finishes when the rat arrives at the goal location at time $t = t_{end}$, and receives an instant reward R.

Figure 5.7 shows the results after 20 trials (for simulation details see Appendix H). The resulting value map, representing the expected reward at any location at the end of the trial, is shown in Figure 5.7a. It is notable that, although the rat ran straight to the goal location, information about the goal location is spread to nearby locations due to the use of the value function approximation and coarse coding.

Figure 5.7b shows the value function (5.32) for all the trials, colour-coded from blue (first trial) to red (last trial). On the first trial (dark blue line)

4 Here Hebbian refers to the conjecture of Donald Hebb. This conjecture states that if input from neuron A often contributes to the firing of neuron B, then the synapses of A to B should be strengthened [4].

the critic does not have knowledge of the goal location, and hence the value function is close to zero. For subsequent trials, the value function increases near to the goal location due to the strengthened connections w^{PC} between the presypnatic activity of the recently visited place cells (before the reward) and the postsynaptic activity of critic cells. After the initial trials, the information about the reward is back-propagated via the eligibility trace. After 20 trials, the value function (dark red line) is close to the correction of the value function $V_c(t)$ in (5.23) (black dotted line).

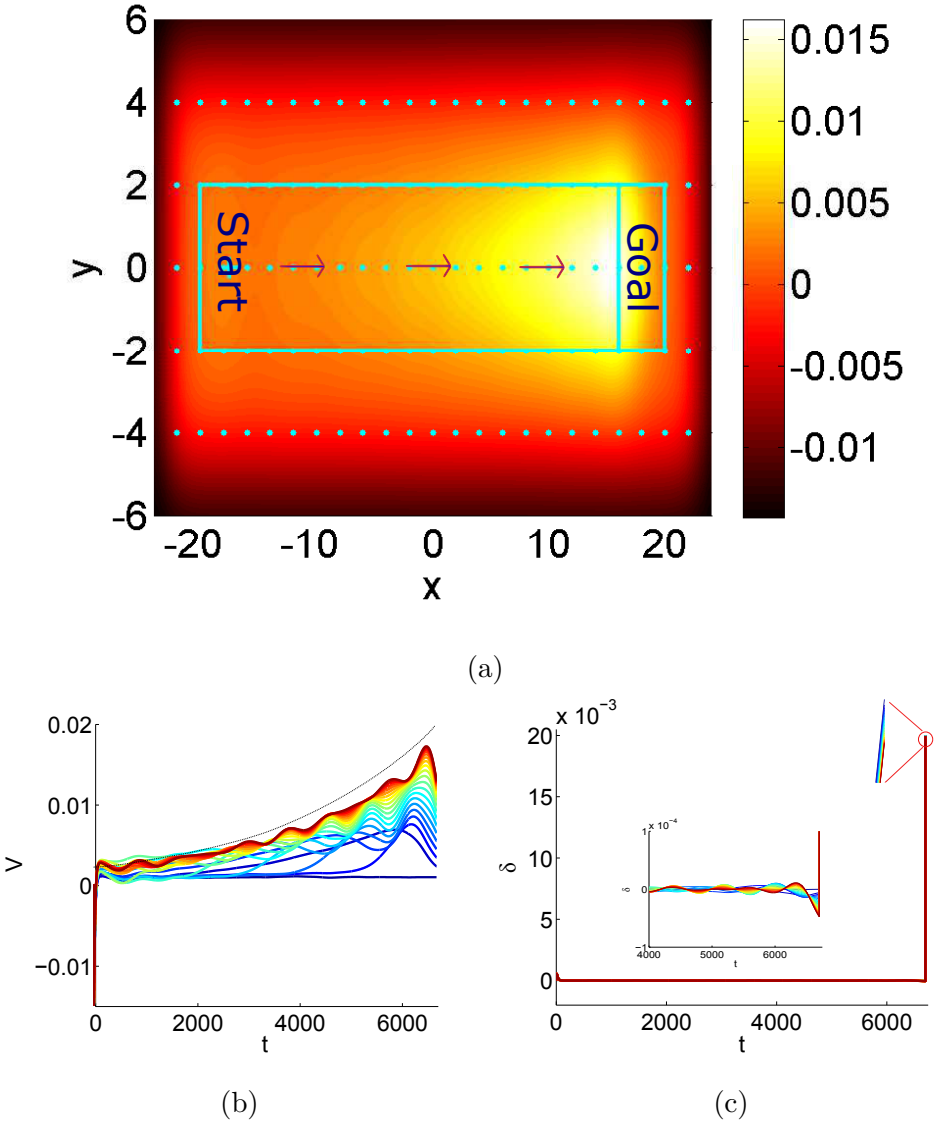
Figure 5.7c shows the TD error given by (5.20) for all the trials colour-coded as above. There is a peak at the end of the trial due to the instantaneous reward obtained at the goal location, which decreases as the agent learns. Smaller changes in the TD error are seen before reaching the goal location due to the back-propagation of information. For example, before reaching the goal location at trial 20 (dark red line) the agent anticipates a reward that is not received at that position, hence δ is negative.

Importantly, the results achieved in this simulation are similar to those obtained in the spiking neuron model proposed by Frémaux *et al.* [88] . Moreover, it suggests that all the modifications to the Frémaux *et al.* model made here are reasonable. These modifications comprise: the change from spike to firing rate based models, the implementation of continuous eligibility traces instead of a "neutral state", and the derivation of the learning rules for the critic. The environment and task description, as well as the architecture remain the same.

5.3.2 *The Actor*

The actor chooses which action to take in order to maximize the reward. Here the actor is modeled as a network of N_A neurons, each coding for a different direction of motion. The activity of the actor evolves according to

$$Q^A u_l^A(t) = \varepsilon_0 \sum_{j=1}^{N_P} w_{lj}^{PA}(t)\rho_j^P(t) + \varepsilon_1 \sum_{k=1}^{N_A} w_{lk}^{AA}(t)u_k^A(t) + \sigma_n n_l(t), \qquad (5.33)$$



(a) (b) (c)

Figure 5.7. Linear Track Simulation. (a) Value Map representing the expected reward at any location in the environment after 20 tirals. Cyan dots represent the centres of place cells and cyan lines delimit the linear track maze. The agent is placed at the start position and its action is forced to go to the right at a constant speed (blue arrows) until it reaches the goal location at $t_{end} = 6700$ ms. (b) Value function as a function of time for each trial (colour-coded from dark blue in trial 1 to dark red in trial 20), black dotted line represents $V_c = \delta(t_{end}) \exp(-(t_{end} - t)/\tau_r), t < t_{end}$ as in Figure 5.5. (c) TD error δ as a function of time. Inset: zoom at times before receving the reward R at the goal location. Colour-coded as in (b). Simulation details are in Appendix H.

with $l = 1, \ldots, N_A$,

$$Q^A = \left(1 + \tau_\vartheta \frac{\mathrm{d}}{\mathrm{d}t}\right)\left(1 + \nu_\vartheta \frac{\mathrm{d}}{\mathrm{d}t}\right), \tag{5.34}$$

and where ε_0 and ε_1 are constants. Here $n(t)$ is a low-pass filtered noise

$$\tau_n \dot{n}_l(t) = -n_l(t) + \Omega_l(t), \tag{5.35}$$

where $\mathbf{\Omega}$ denotes normal Gaussian noise. The size of the perturbation σ_n is modulated by the performance, such that the noise decreases as the performance increases. Hence

$$\sigma_n = \sigma_{n0} \min\left[1, \max\left[0, \left(V_{max} - V(t)\right) / \left(V_{max} - V_{min}\right)\right]\right], \tag{5.36}$$

where V_{min} is the minimal level of expected reward, whereas V_{max} is the maximal. Here $\sigma_{n0} > 0$ is a constant. The first term on the right hand side of (5.33) represents the connections between place cells and actor cells (analogous to the corresponding term in (5.30)). The next term in (5.33) gives the lateral feedback of actor cells, the connectivity $w^{AA} \in \mathbb{R}^{N_A \times N_A}$ excites neurons nearby and inhibits all other neurons. This term promotes the formation of a single bump of activity. During early trials the bump will be formed at a random position (driven by noise), but on learning the appearance of the bump becomes more deterministic (due to the strengthening of synaptic weights between place and actor cells and the decrease of the noise perturbation σ_n in (5.33)). This results in a good exploration/exploitation balance. The activity of the actor differs from that in [88] in that a noise term is added explicitly. Note that in the model proposed by Frémaux *et al.* there is a form of noise due to the spiking activity whereas here it is added in order to promote exploration.

The agent's choice is to go towards the direction coded by the actor cell with maximum activity. Let θ_l be the direction coded by the lth actor cell, then the state evolution is given by

$$\dot{\mathbf{x}}(t) = s\begin{pmatrix}\cos(\theta^*)\\ \sin(\theta^*)\end{pmatrix}, \qquad \theta^* = \theta\left(\max\left[\rho^A(u^A; w^{PA})\right]\right), \tag{5.37}$$

where s is a constant speed and $\rho^A = \rho^C$ as given by (5.31). Here, the update of the weights between presynaptic place cells and postsynaptic actor cells is given by

$$\dot{w}_{lj}^{PA} = \eta_A \delta(t) \frac{\partial \rho^A(u_l^A; w^{PA})}{\partial w_{lj}^{PA}}, \tag{5.38}$$

where η_A is the constant learning rate of the actor and $\delta(t)$ is the TD error given by (5.20). With this learning rule the connection between neurons forming the bump activity of actor cells and the recently visited place cells will be strengthened when $\delta(t) > 0$ and weakened when $\delta(t) < 0$. Note that the bump of activity formed by the connectivity of actor cells and the chosen ρ^A[5] ensures generalisation across the actor neurons because actions coding for similar direction will be reinforced according of the performance of the action.

5.3.3 *Simulation Results*

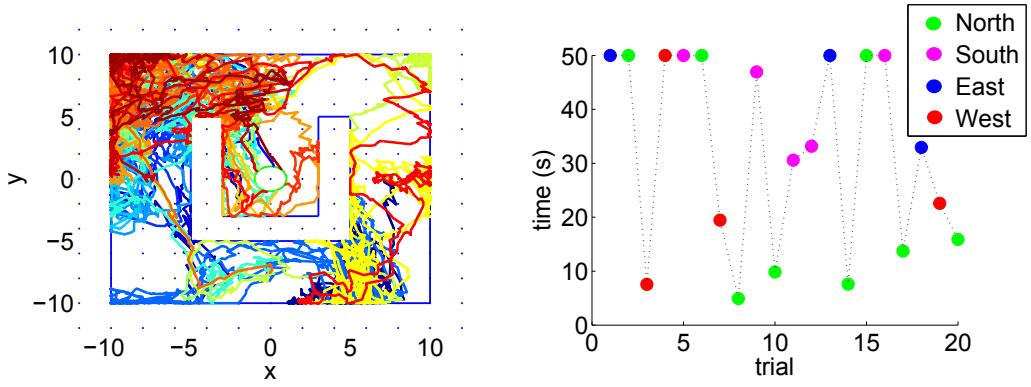
In order to confirm that the rate network model for the actor and critic dynamics proposed in this work can achieve similar results to those in Fremaux *et al.* [88] (Fig. 5.6) 20 trials were simulated. On each trial the agent starts at one of 4 randomly-chosen starting positions. The maze is a square environment with a circular goal location at the centre surrounded by a U-shaped obstacle, as shown in Figure 5.8. The trial ends when the agent reaches the goal location and gets an instantaneous reward R_{goal} or when $t > t_{max}$, at which point the agent is removed from the maze. All the agent's trajectories are shown in Figure 5.8a, and exhibit a good exploration/exploitation balance, covering a large proportion of the maze. Figure 5.8b shows the time taken by the agent to find the goal location at each trial. It can be seen that at early trials the agent was not able to find the goal location from any starting location, except on the third trial when the agent found the goal location starting at the western starting position. After the sixth trial, the agent reached the goal location in

5 Note that ρ^A is an exponential function, so the change in w^{PA} will be proportional to its derivative. See Appendix I for details.

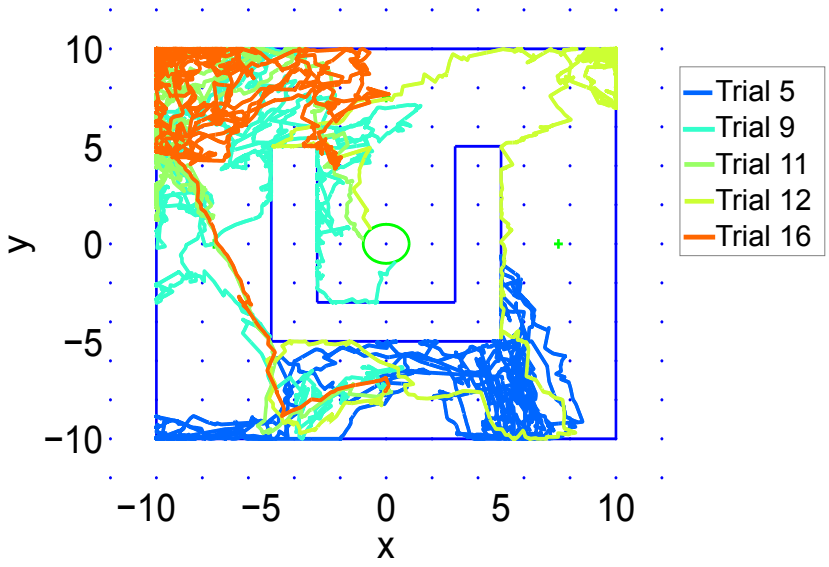
most of the trials, starting at any of the initial positions. It was easier for the agent to find the goal location when it started at the northern location due to the obstacle.

Figure 5.8c illustrates how the trajectories starting from the southern location change as the agent learns the goal location. The first time the agent started at the southern start location was on the fifth trial (blue line), the agent was not able to find the goal location and was removed from the maze after t_{max} ms. The following occasion was on the ninth trial (aqua line). This time, instead of staying in the same area as in the fifth trial, the agent moved to the western part of the maze, and when it approached the goal location it remained close to it. This is because the agent was able to exploit information about the goal location, via its value function, obtained from western starting positions on trials 3 and 7, and from north starting position on trial 8. At trial 11 (light green) the artificial rat followed a similar path as in trial 9, but avoided the interior wall of the U-shaped obstacle. On the next trial (12, lime green), the agent took the same route as in trials 9 and 11, but then it chose to go right instead. Note that since the southern starting position is far from the goal location (due to the obstacle), the information about the goal location had not yet back-propagated sufficiently; allowing the agent to explore other options. Eventually, however, it reached the north part of the maze where it found the goal location. On trial 16 the agent was not able to reach the goal location. This could have been due to the fact that the synaptic weights between the place and actor cells decreased in locations at the top left corner of the maze, allowing the noise term to take control, which translated to randomly chosen actions.

Figure 5.9 shows the value maps for this task on different trials. Initially, synaptic weights between the place and critic cells were randomly selected (Fig. 5.9a). After seven trials (Fig. 5.9b) these weights decreased near the maze and obstacle walls, due to the negative reward received when a collision occurred. The value map has its highest value in the unexplored top right corner. In Figure 5.9c the value map after trial 14 has low values even near the goal



(a) (b)



(c)

Figure 5.8. (a) Colour-coded trajectories from an agent during 20 simulated trials of the maze task. Here earlier trials are represented by dark blue lines, and later trials by dark red lines. The goal location (green cicle) is surrounded by a U-shaped obstacle. Starting positions are represented by green crosses equidistant from the goal location. Blue dots denote the place cell firing fields centres. (b) Time in seconds taken by the agent to find the goal location at each trial, starting at the North (green), South (magenta), East (blue) and West (red) starting postions. (c) Trajectories starting at southern starting location, $\mathbf{x} = (0, -7.5)$, corresponding to the trial numbers 5, 9, 11, 12 and 16. Simulation details and parameters are presented in Appendix I.

location due to collisions with the obstacle's internal walls. After all the trials the value map (Fig. 5.9c) has a bump of activity on the north part of the maze, near the goal location due to the back-propagation of the information about the goal location. Note that near the goal location the value function decreases due to the negative reward from colliding with the walls. Other bumps of activity are seen in unexplored areas.

Figure 5.10 shows the policy map (how likely it is that the agent chooses a certain direction). Because the initial synaptic weights between place and actor cells were randomly chosen, the initial policy map (Fig. 5.10a) is also random. After 7 trials (Fig. 5.10b) it can be seen that the policy dictated directions around the top left corner of the maze are weak, and hence the associated actions are effectively random. Note that there is also a change of preferred direction around the goal location. For later trials (Figs. 5.10c and 5.10d) the policy map changes due to collision with the walls, rather than from the learning of the goal location.

These results show that the proposed firing rate description of the actor activity and action selection (and therefore the state evolution) are able to achieve good performance in the task proposed by Frémaux *et al.* [88]. Simulations where only the actor activity was changed from the one in the spiking model show the same behaviour (data not shown), validating the firing rate description of the actor activity proposed here.

The rate description of the critic and actor activity proposed in this section embedded in the continous time framework proposed by Frémaux *et al.* [88] showed good performance. The agent was able to reach the goal location in less time as it explored (and learned using reinforcement learning) the environment.

First, the critic component of the model was studied. The spiking activity was transformed to a rate description. Furthermore, the implementation of eligibility traces for the learning rules, instead of a "neutral state", has a more solid theoretical basis. It was shown that after 20 trials the value function converges to the theoretical value.

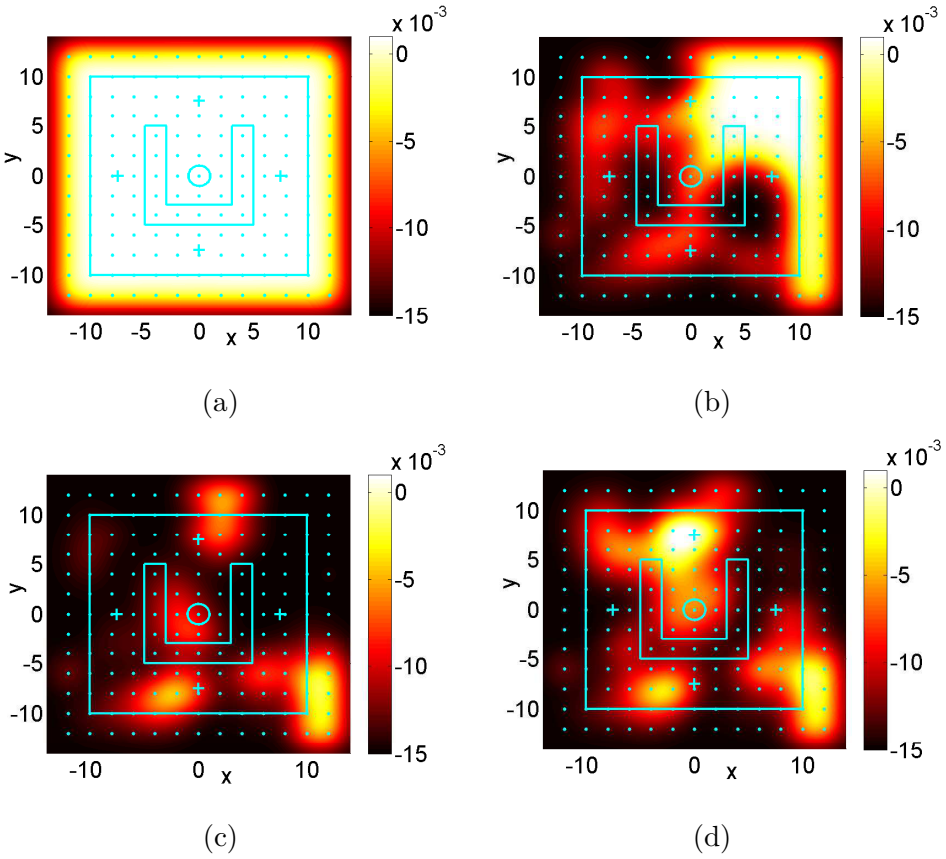


Figure 5.9. Value map (colour map) representing the learning of the place-critic cell synaptic weights by the agent on the maze task in Figure 5.8 before the beginning of the task (a), and after trials 7 (b), 15 (c), and the last trial (d). Simulation details and parameters are presented in Appendix I.

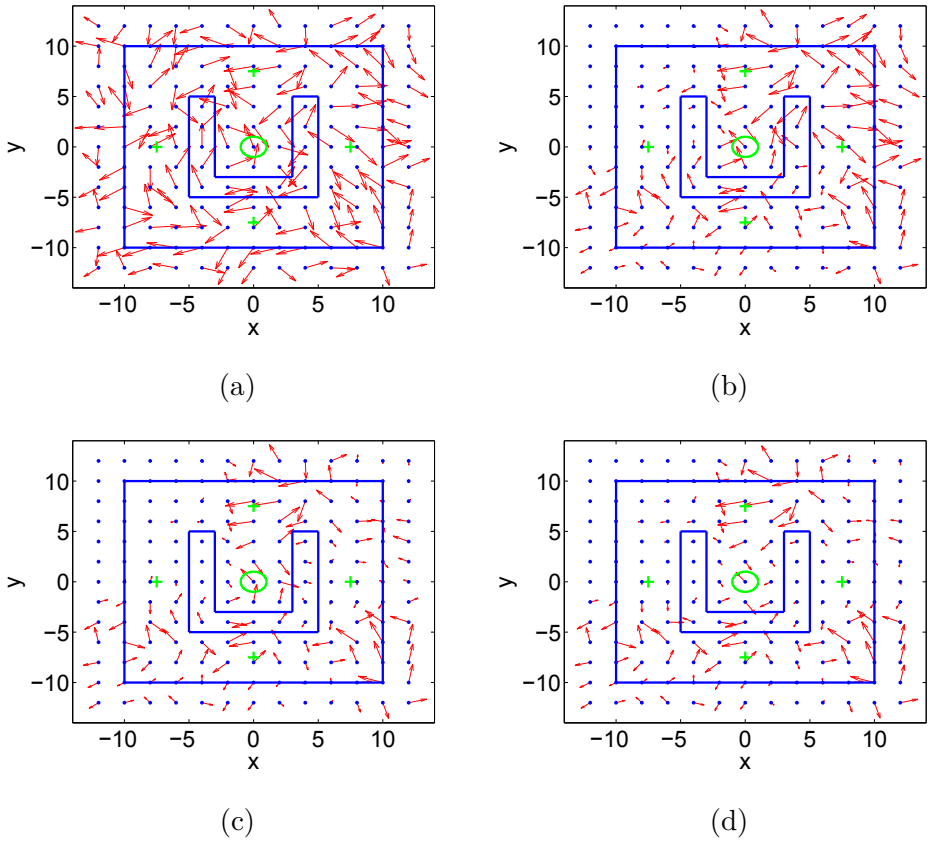


Figure 5.10. Policy map (vector field) representing the learning of the place-actor cell synaptic weights by the agent on the maze task in Figure 5.8 before the beginning of the task (a), and after trials 7 (b), 15 (c), and last trial (d). Simulation details and parameters are presented in Appendix I.

As with the critic, rate activity was considered for the actor. Additionally, a noise component was included to the actor activity in order to promote exploration. This was modulated by the performance, so if the agent was near to the goal location the noise was decreased. However, it was noted that when the agent collided with the walls and received a negative reward, the synaptic weights between the place and critic/actor cells decreased to a point that noise dominated and actions were taken randomly. This randomness is necessary to avoid further collision with the walls, but in later trials it in fact prevented the artificial rat from reaching the goal location.

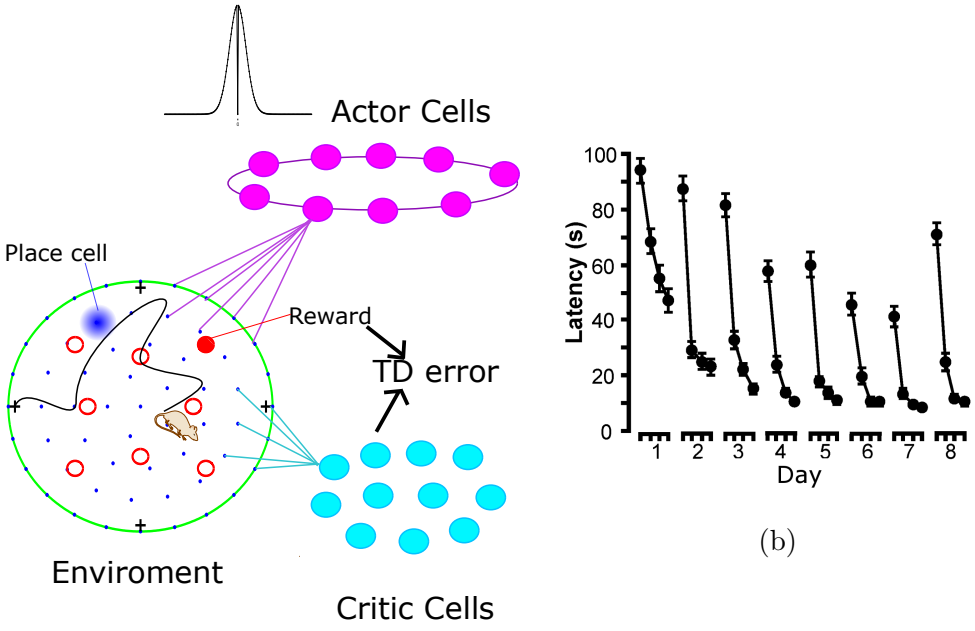
A further modification to [88] is that the preferred direction was chosen based on the direction coded by the actor neuron with highest activity, instead of multiplying the activity of each actor neuron by its preferred direction. This methodology gives similar results in simulations and was chosen because in the next section the agent is forced to go in a certain direction towards the goal location, if it is unable to reach it within a given time; hence the agent either chooses the direction coded by the actor cell with higher activity or the forced direction. Both action selection choices gave similar results for this task.

The learning rules for the synaptic weights between the place and actor cells were updated using the TD error and the change of the actor firing rate with respect to the weights. Eligibility traces for the actor were not used because the bump of activity formed on the actor slowly changes location over time, allowing the actor to have some memory of the actions taken on previous states (the weights between actor and place cells are updated proportionally to the position and amplitude of this bump). Importantly, it was observed that the agent learns how to avoid collision with walls instead of reaching the goal location because the received negative reward has a stronger learning impact than the positive reward at the goal location.

5.4 DELAYED-MATCHING-TO-PLACE WATERMAZE TASK

In this section the proposed model with rate firing activity for the critic and actor is implemented on a delayed-matching-to-place (DMP) watermaze task, motivated by the experiments in Bast *et al.* in [90]. The DMP water maze task, illustrated in Figure 5.11a, is a variation of the Morris water maze task. In both tasks, a rat is placed into the so-called watermaze, a circular tank filled with opaque water and surrounded by spatial cues, that contains a hidden escape platform. Because swimming fatigues the rat, it is highly motivated to find the platform. In the DMP task, the platform location changes each day. The DMP task spans 8 days with 4 trials each day. Each trial starts from one of 4 possible initial positions chosen at random. Interestingly, in experiments rats show *one trial learning*, whereby the the latency and pathlength travelled to find the platform steeply declines from trial 1 to 2, with little further reductions on the subsequent trials, as shown in Figure 5.11. There is also evidence that rats improve their overall efficiency on the task, with latencies and pathlengths during trial 1 tending to reduce across days.

Foster *et al.* [79] presented a model that simulates the DMP task using reinforcement learning in an actor-critic discrete time framework. In their model, the critic is represented by a single cell, whereas the actor is represented by a population of 8 independent neurons, each coding for a different direction. The actor and the critic receive information about the environment via place cells. The actions are chosen stochastically. However, this simple structure is unable to solve the DMP task, and therefore Foster *et al.* added a second component to their model: a network that learns spatial coordinates, using inputs from place cells. This component computes which direction the agent needs to swim in order to reach the goal location from its current position. This direction is integrated to the actor-critic component by adding an action that represents the agent's preference for swimming in the direction offered by this coordinate system.



(a) (b)

Figure 5.11. (a) Delayed-matching-to-place watermaze task. Left: maze environment. Blue dots represent the centres of place cell firing fields. Red circles represent the goal locations. Blue shadow shows how activity of a place cell decays as the distance from the centre of its firing field increases (5.29). Black line: trajectory of the agent from the starting position (black croses). When the agent reaches the goal location (filled red circle) it receives a reward. Top right: actor neurons. Bottom right: critic neurons. (b) Latencies (mean ± SEM-standard error of mean) to reach the platform are plotted for the 4 different trials each day, across 8 days, for the DMP task. Figure modified from [90], 89 rats were used in this experiment.

The aim of implementing the rate model on a DMP task is to test if the agent can learn to find to the goal location in an environment taken from real experiments, without the presence of obstacles. Moreover, it is of interest to determine if the proposed model is capable of one trial learning. Such a model will allow the exploration of hypotheses of the effects of hippocampal lesions in behaviour (as those in [90]). This could be done by representing the lesions as perturbations in the connectivity between place cells and critic/actor cells, and assessing the behaviour as the agent's performance during the task.

5.4.1 *Simulation Results*

The actor and critic evolution is simulated as in §5.3. In order to avoid the slow learning due to the agent colliding with the walls that was observed in the previous task, it is assumed that the artificial rat prefers to swim towards the maze centre when it is near the walls. This is implemented by modifying the initial synaptic weights between the place and actor cells at the maze wall accordingly. This modification is realistic since the rats learn to avoid the walls and swim towards the goal location in "pre-training" sessions. Previous place cell models, as in [84], have implemented boundary-related information; furthermore, there is experimental evidence that a type of neuron, so-called border cells, fire at the walls or boundaries of an enclosure [91].

The trial ends when the agent reaches the selected hidden platform, but unlike the previous task, when $t > t_{max}$ the agent is guided towards the goal location by forcing it to take a specific action until it reaches the goal location and obtains a reward, as in experiments [90]. Note that the size of the maze is larger than in the previous study (the area of the maze in §5.3 is 400 cm^2 and the area of the watermaze in this section is ≈ 31416 cm^2), so the place cell firing function parameters are changed as a consequence. The place cell firing field centres are arranged in concentric circles. Simulation details can be found in Appendix J.

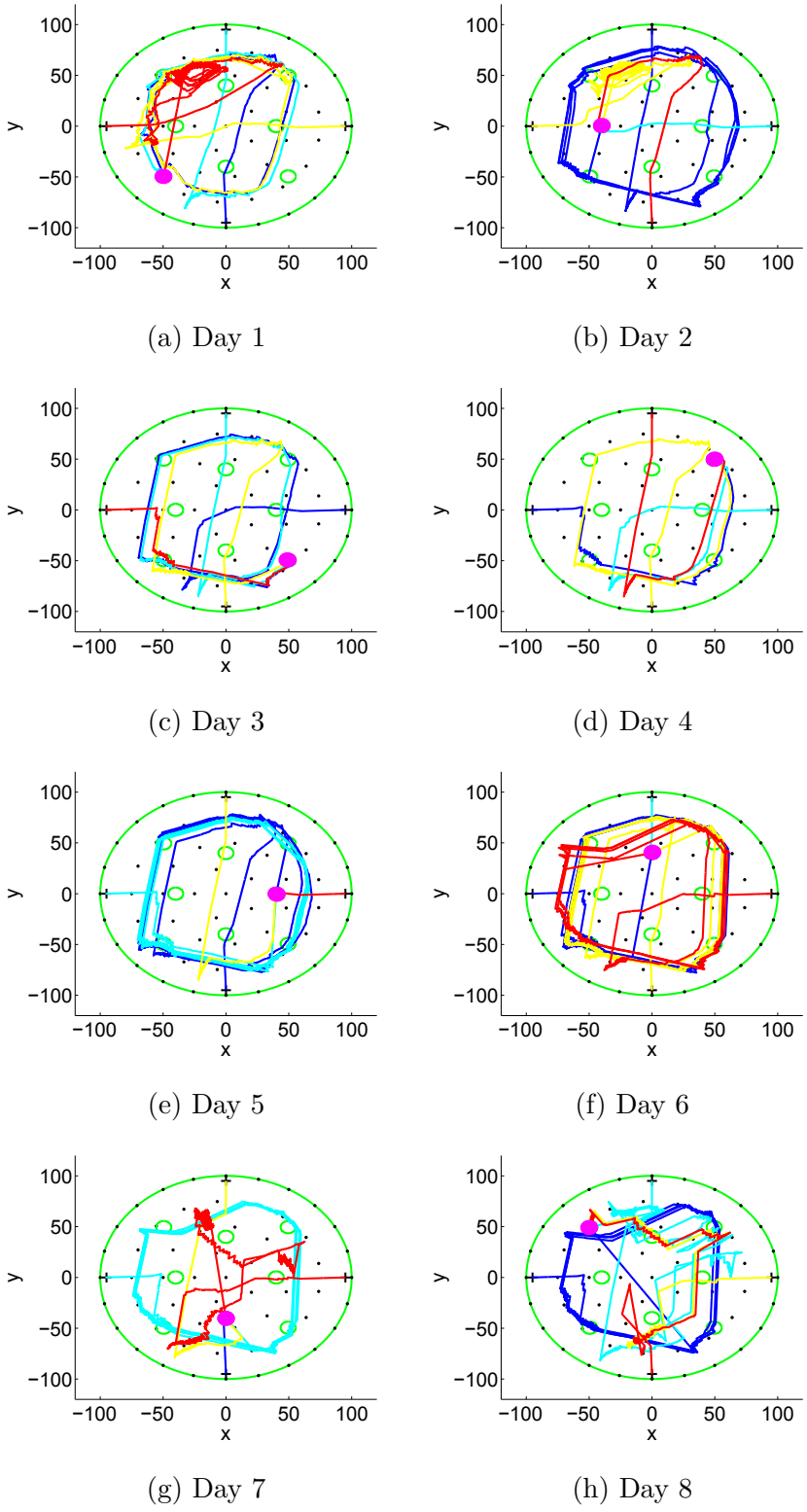
The agent trajectories are shown in Figure 5.12. Note that at the start of a trial the agent swims towards the maze centre, and, moreover, it avoids collision with the wall during the experiment, due to the initial modification to the weights described above. Figure 5.13 shows the changes in the value function during the experiment. The peak of activity after the fourth trial every day is located near to the selected goal location, showing that the critic is able to learn the new location each day. The change of the peak location between days 2 and 3 can be seen in Figure 5.14, after the first trial on day 3 (Fig. 5.14e) there are 2 peaks of activity, one corresponding to the goal location from the previous day and the other signaling the new goal location. After the third trial on day 3, the peak of activity corresponding to the previous goal location vanishes due to the negative TD error computed when the agent does not receive the expected reward at that location.

On the other hand, the actor performance is poor. As shown in Figure 5.15, although the policy map is able to change towards the goal direction, such changes do not back propagate far enough to enable the agent to take shorter paths towards the goal. This could be due the difficulty of gaining knowledge of the value of all the possible actions across the maze.

Figure 5.16 shows the latency time of this experiment. The agent was able to find the hidden platform in 20 trials, but it had to be led to the goal on 12 trials (accounting for more than one third of the trials). There is no one trial learning in Figure 5.16. It is noticeable that on days 5 and 8 the agent showed an improvement by reducing steeply the latency within two trials.

5.4.2 *Visual Cues*

As mentioned above, the proposed firing rate critic-actor framework is incapable of achieving one trial learning as in experiments done in [90]; therefore, additional mechanisms are included in order to improve the performance. For example, visual cues are available outside the watermaze during the experiments and are believed to play a key role in the rat's ability to orientate itself.



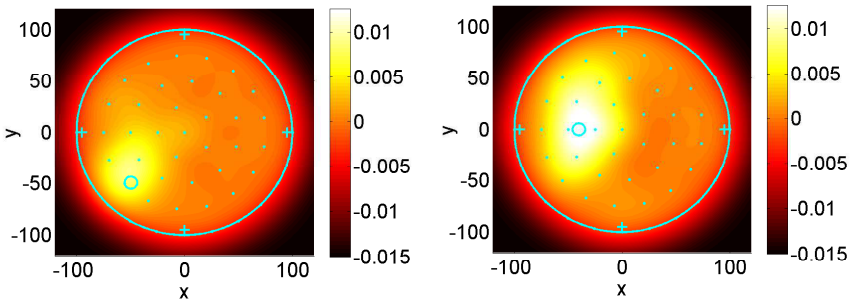
(a) Day 1 (b) Day 2

(c) Day 3 (d) Day 4

(e) Day 5 (f) Day 6

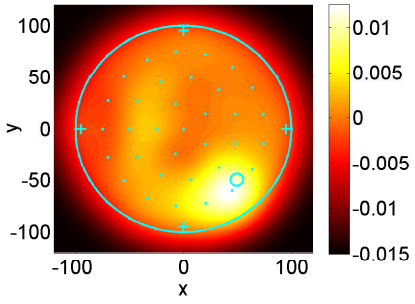
(g) Day 7 (h) Day 8

Figure 5.12. Colour-coded trajectories for 4 trials on the DMP task for different days. Here trajectories of trials 1,2,3,4 are represented by the blue, cyan, yellow and red line respectively. Black dots are the centres of place cell firing fields and black crosses are the starting locations. The magenta circle represents the chosen escape platform for a given day, and green small circles are the goal locations for the other days. Simulation details and parameters are presented in Appendix J.

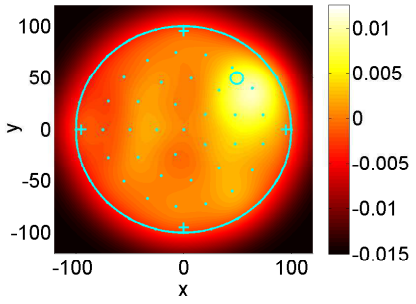


(a) Day 1

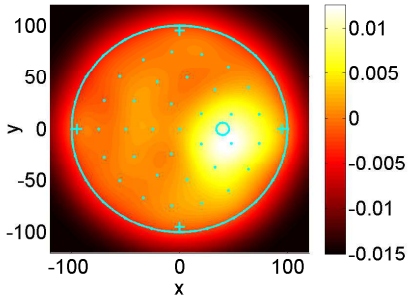
(b) Day 2



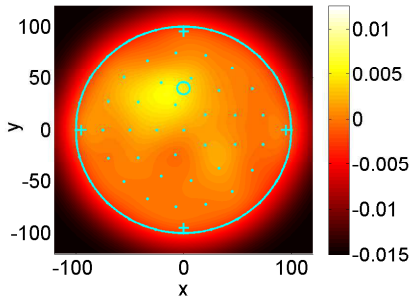
(c) Day 3



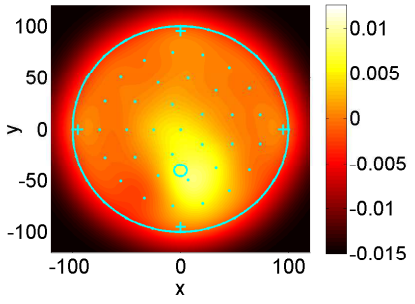
(d) Day 4



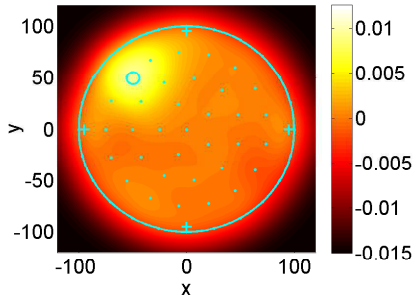
(e) Day 5



(f) Day 6

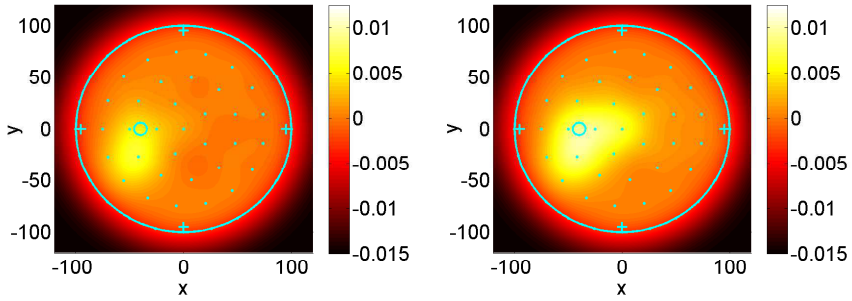


(g) Day 7



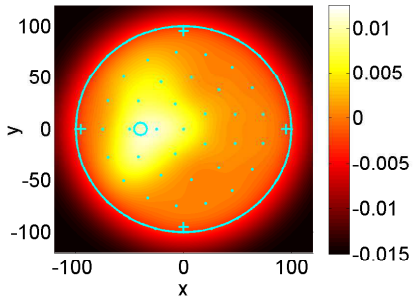
(h) Day 8

Figure 5.13. Value map of the DMP task in Figure 5.12 after 4 trials for each day. Simulation details and parameters are presented in Appendix J.

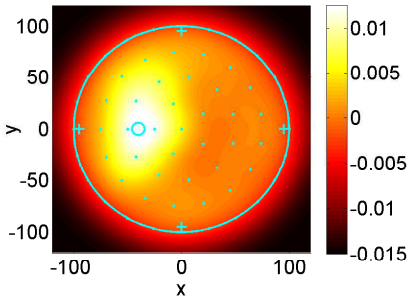


(a) Day 2, Trial 1

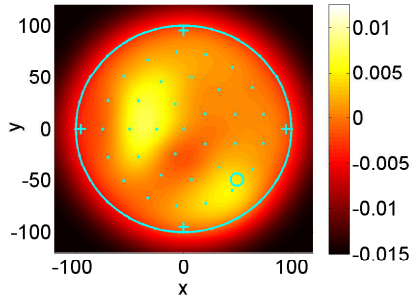
(b) Day 2, Trial 2



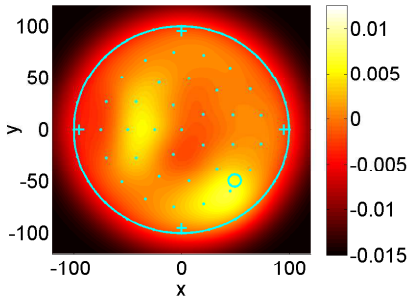
(c) Day 2, Trial 3



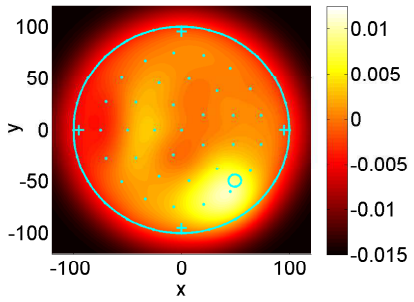
(d) Day 2, Trial 4



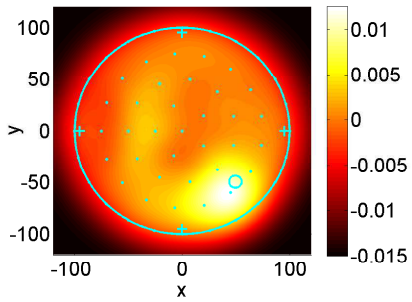
(e) Day 3, Trial 1



(f) Day 3, Trial 2

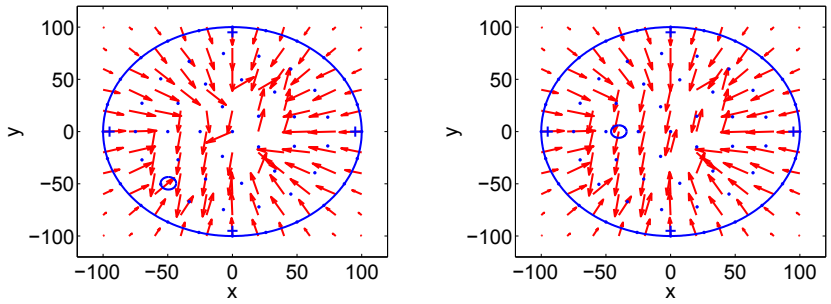


(g) Day 3, Trial 3



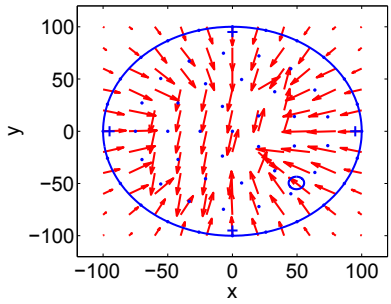
(h) Day 3, Trial 4

Figure 5.14. Evolution of the value map of the DMP task in Figure 5.13 after each of the 4 trials for days 2 and 3, showing how the peak of activity changes location from the past (day 2) to the new (day 3) goal location. Here the selected goal locations are represented by cyan circles. Simulation details and parameters are presented in Appendix J.

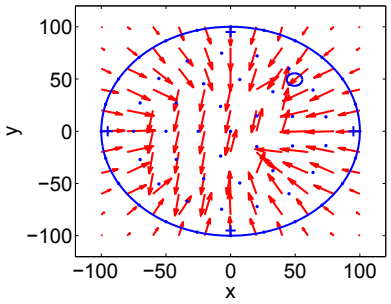


(a) Day 1

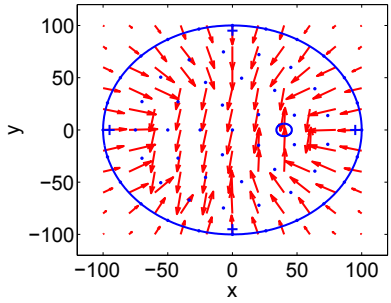
(b) Day 2



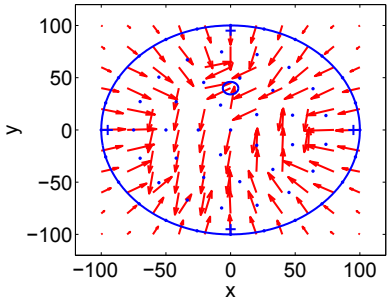
(c) Day 3



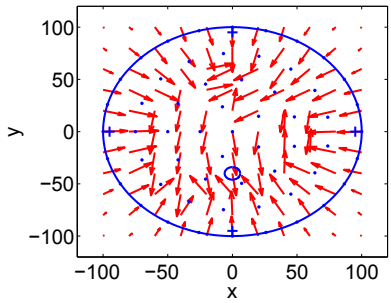
(d) Day 4



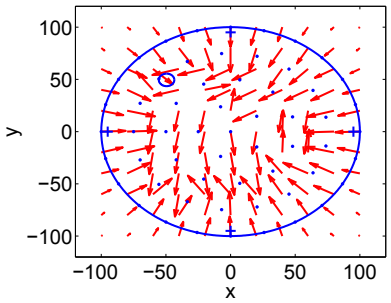
(e) Day 5



(f) Day 6



(g) Day 7



(h) Day 8

Figure 5.15. Policy map of the DMP task in Figure 5.12 after 4 trials for each day. Simulation details and parameters are presented in Appendix J.

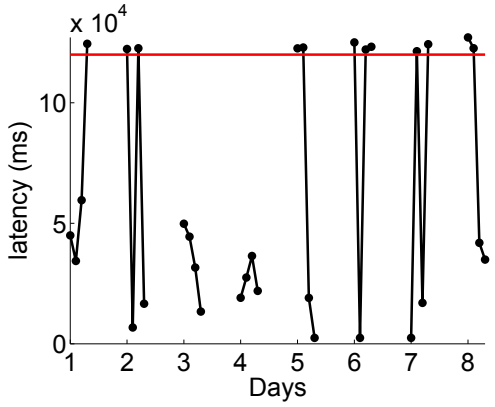


Figure 5.16. Latencies to reach the platform are plotted for the 4 different trails each day (dots) across 8 days (lines) of the experiment in Figure 5.12. Red line represents $t = t_{max}$.

Therefore, 8 visual cues are added to the environment and the algorithm is modified in the following way:

- On the first trial of each day the simulation is performed without any visual input. When the agent finds the hidden platform it is assumed that the agent relates the three nearby cues to the found goal location (see Fig. 5.17). One of these visual cues corresponds to the nearest cue from the goal location and the other two visual cues are adjacent to the former one.

- During trials 2, 3 and 4 the agent uses the nearby cues to check its relative position to them every 10 s, for 2.5 s or less. During this period the agent first checks if it is near the goal location (yellow zone in Fig. 5.17). If the agent is near the goal location, the checking ends. If not, the agent checks if it is facing any of the selected visual cues (purple asterisks in Fig. 5.17), and if so the checking ends. If not, the agent receives a punishment with strength $R_{obst}/5$ and is forced to go in the opposite direction with a step length of $|\delta(t)|$.

Figure 5.18 shows the agent's trajectories during an experiment. Note that on days 1 to 3 the performance is better than in later days, and moreover this performance is better in comparison with the experiment without visual cues.

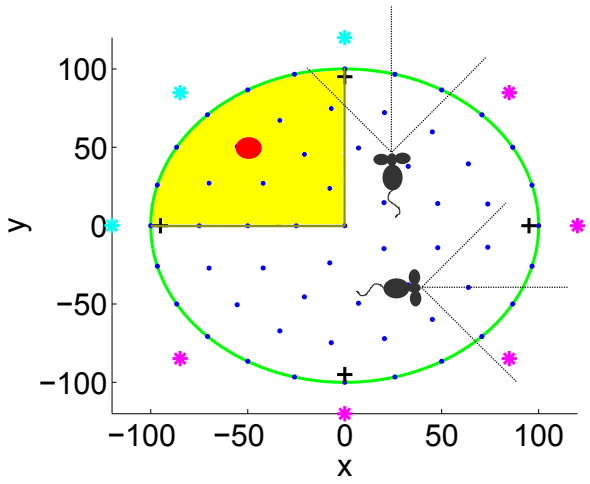


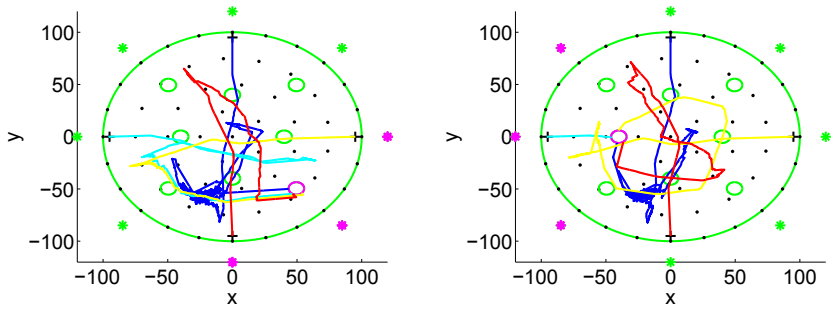
Figure 5.17. Navigation using visual cues. Blue dots: place cell firing field centres. Red disk: escape platform. Magenta stars: visual cues outside the watermaze; cyan stars: selected cues nearby the goal location. The yellow zone represents the zone near the escape platform. There are two synthetic rats in this maze illustrating different locations of an agent during a trial, and how these locations relate to the goal location. The agent at the top is not in the zone near to the goal location (i.e., it is outside the yellow zone) but is facing one of the selected visual cues, therefore no punishment is applied. The synthetic rat at the bottom is outside the yellow zone and is not facing any of the selected cues, thus it is forced to change direction and receives an instantaneous negative reward.

The learning by the critic is represented by the value maps in Figure 5.19. As in the experiment without visual cues, the critic is able to learn the location of the selected escape platform each day; however, unlike that experiment, there are some locations inside the watermaze that have more negative values due to the punishment received when the agent was in a poor position relative to the goal location and nearby cues.

Importantly, the addition of visual cues to the simulation translates to an improvement in learning the optimal choice of direction by the actor, as can be seen in the policy map shown in Figure 5.20. On the other hand (with the exception of days 1-3 as noted above), if the overall performance of the agent is measured as the reduction of the latency time after the first trial every day and across the days (see Fig. 5.11), then the experiment with visual cues (Fig. 5.21) actually shows poorer performance than that without (Fig. 5.16). In fact, in the experiment with visual cues the agent had to be guided to the goal location more times than in the one without visual cues.

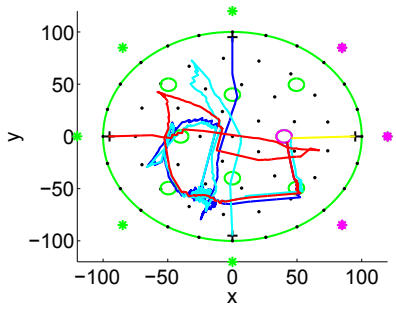
In order to reduce any bias due to a specific set of platform positions across days and starting positions across the experiments, the simulation with and without visual cues were run with the same escape platforms and starting positions (described in Appendix K). Also, to avoid bias due to a specific selection of initial weights between the presynaptic place cells and postsynaptic actor and critic cells, the same experiment was performed with 15 synthetic rats with different initial weights. Figure 5.22 shows the resulting latency (mean $\pm$ SEM) for the experiment with (Fig. 5.22b) and without (Fig. 5.22a) visual cues. The use of visual cues improves the latency on days 4 and 6, but on early days (1-3) the agents not using visual cues had better performance. Also, on day 5 and 7 the group without visual cues reached the escape platform in less time on average.

In conclusion, the implementation of visual cues shows a better learning by the actor, but this does not translate to an improvement of the overall performance of the task. Changes to the position checking (relative to the visual cues) frequency and duration does not improve the performance. Actually, when the

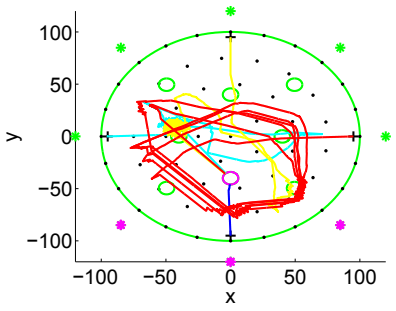


(a) Day 1

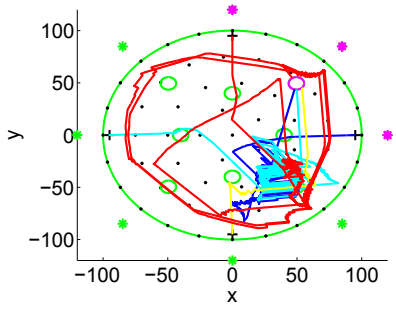
(b) Day 2



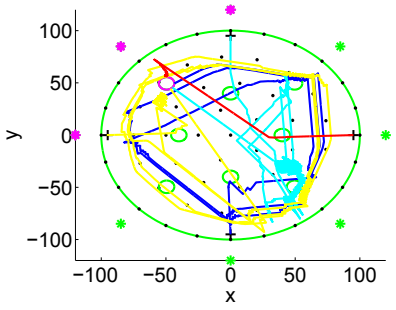
(c) Day 3



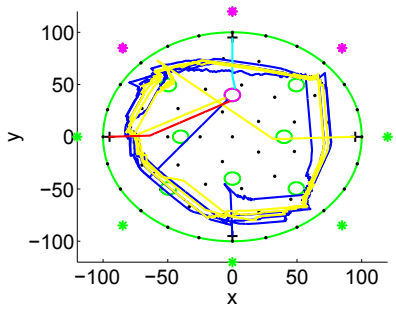
(d) Day 4



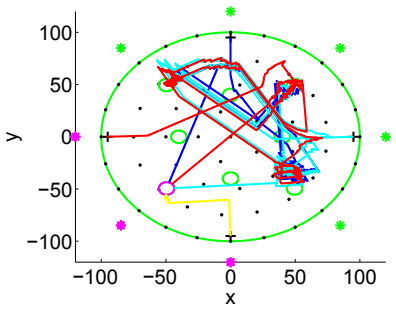
(e) Day 5



(f) Day 6

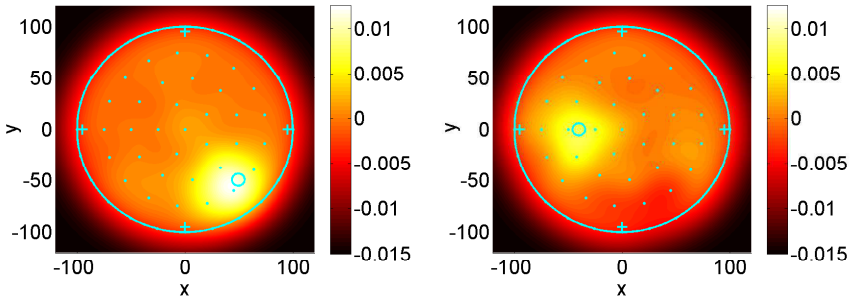


(g) Day 7



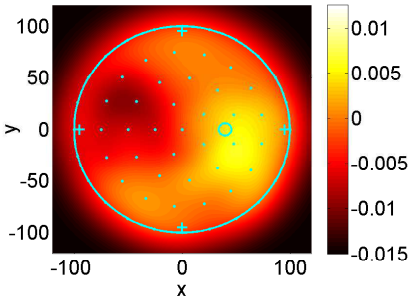
(h) Day 8

Figure 5.18. Colour-coded trajectories for 4 trials, from blue (trial 1) to red (trial 4), on the DMP task with visual cues for each day. Black dots: centres of place cells firing fields, black crosses: starting locations. Magenta circle represents the chosen escape platform whereas green small circles are the other platforms. Asteriks: visual cues, the magenta ones represent the cues nearby the goal location. Simulation details and parameters are presented in Appendix J.

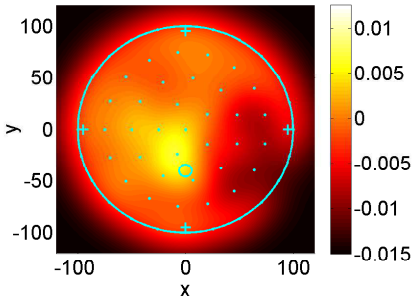


(a) Day 1

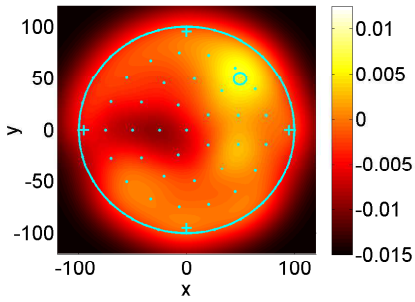
(b) Day 2



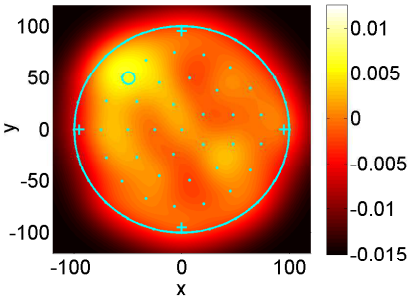
(c) Day 3



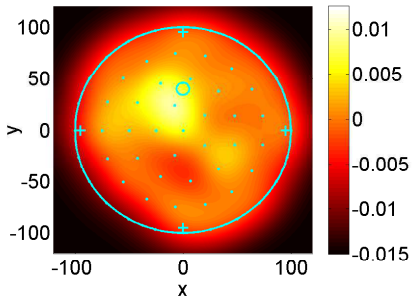
(d) Day 4



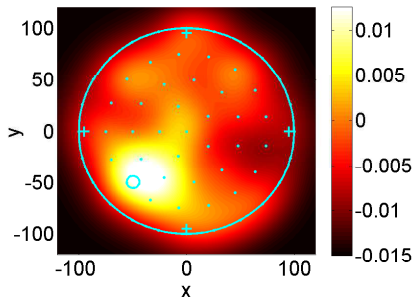
(e) Day 5



(f) Day 6

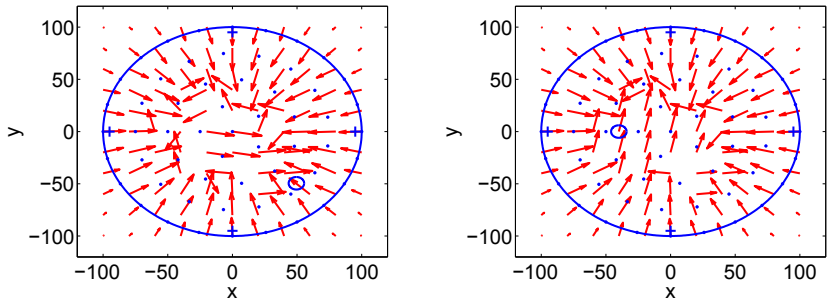


(g) Day 7



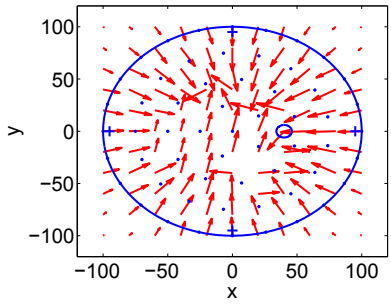
(h) Day 8

Figure 5.19. Value map of the DMP task with visual cues in Figure 5.18 after 4 trials for each day. Simulation details and parameters are presented in Appendix J.

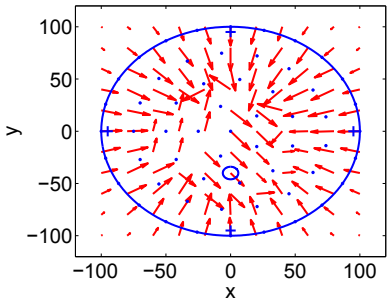


(a) Day 1

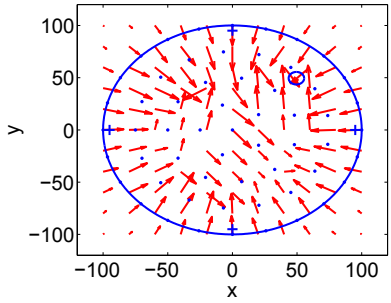
(b) Day 2



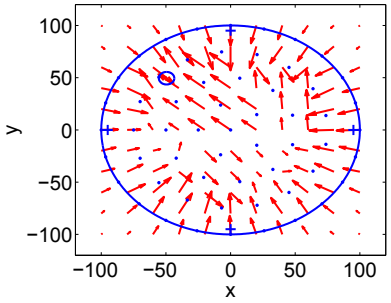
(c) Day 3



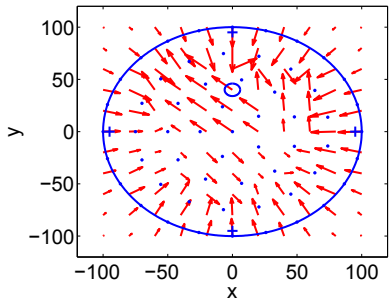
(d) Day 4



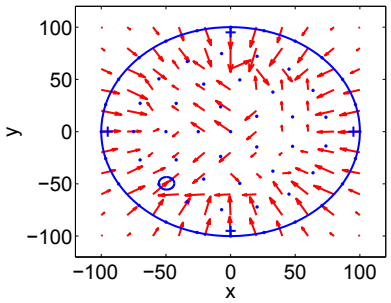
(e) Day 5



(f) Day 6



(g) Day 7



(h) Day 8

Figure 5.20. Policy map of the DMP task with visual cues in Figure 5.18 after 4 trials for each day. Simulation details and parameters are presented in Appendix J.

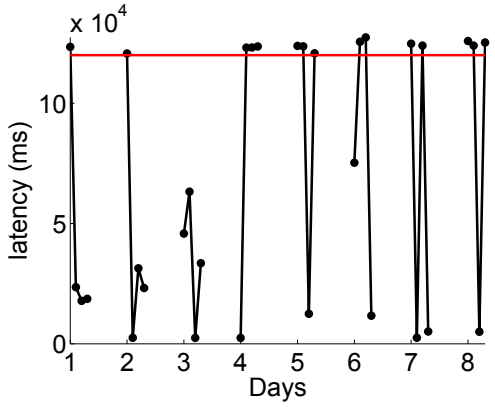


Figure 5.21. Latencies to reach the platform are plotted for the 4 different trails each day (dots) across 8 days (lines) of the experiment with visual cues in Figure 5.18. Red line represents $t = t_{max}$.

frequency and/or the duration of the checking increases, the weights between place cells and actor neurons decreases, due to the punishment received when the agent is not in the zone near the goal location or facing a selected visual cue. Thus, making the agent more likely to make random choices. Additionally, weights between place cells and critic neurons decrease, and sometimes this leads to a disruption of learning by the critic. Importantly, a decrease of the punishment when the agent is far from the goal location and not facing toward any of the visual cues decreases the performance of the actor.

5.4.3 *Increase in Size and Spacing of Place Cell Firing Fields*

In this subsection the effect of increasing the size and spacing of place cell firing fields is investigated. Here, the centres of the place cell firing fields are distributed on a square grid, similar to the grid used for the maze task in §5.3, instead of concentric circles. Visual cues are not implemented. As shown in Figure 5.23, the new distribution of place cell firing fields centres allows the agent to trace more direct paths towards the goal location is some cases. Although the number of place cell firing fields decreased from 58 to 25, the critic shows good performance, as illustrated in Figure 5.24. The policy map in Figure 5.25 shows a performance similar to that on Figure 5.15.

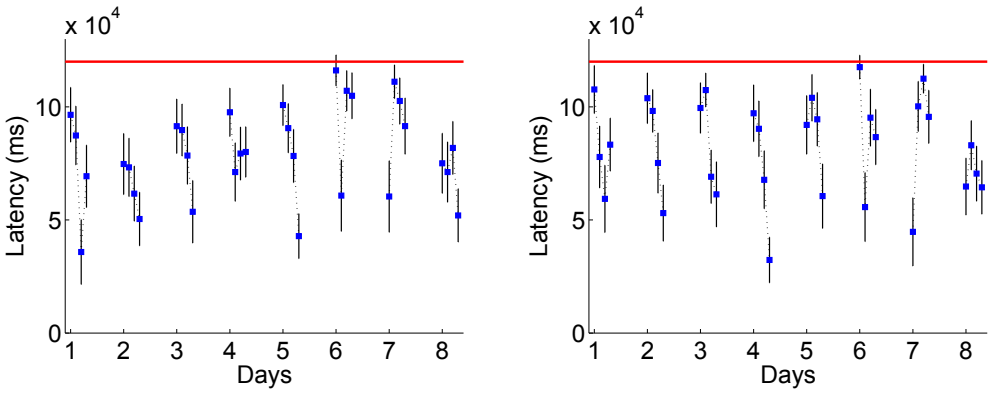
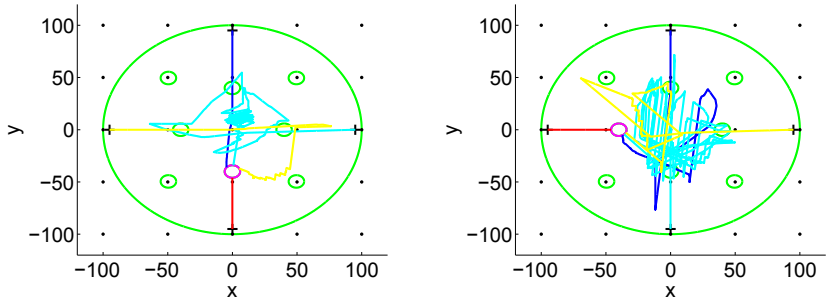


Figure 5.22. Latency mean (blue squares) ± SEM (black lines) for 4 trials across 8 days for the DMP task. Maze with (b) and without (a) visual cues. Here 15 artificial rats were simulated using the protocol in Appendix K

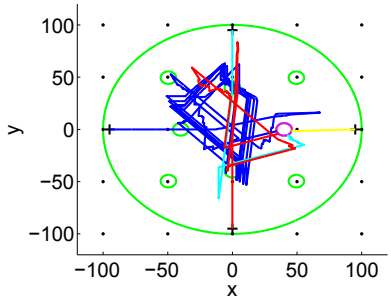
Although the latencies to reach the platform are similar for both firing fields distributions for a particular agent (Figs. 5.16 and 5.26a), important differences were found when comparing the latency mean for 15 artificial rats with the experimental protocol in Appendix K (Figs. 5.22a and 5.26b). For example, on the first trial of the first day, the agents spend more time (on average) trying to find the goal location in the maze with larger firing fields, but these agents were more effective on the later trials of that day. In general, there is a reduction of the latency for the DMP task using the maze with larger place cell firing fields, although on the first trial on days 2 and 4 all the agents were unable to find the goal location and had to be guided.

The agents were not capable of one-trial learning on the DMP task with an increase in size and spacing of place cell firing fields, however, they showed an important decreased in the mean latency time to find the goal location. This results suggest that the distribution of the place cell firing field are crucial to the performance of the agent.

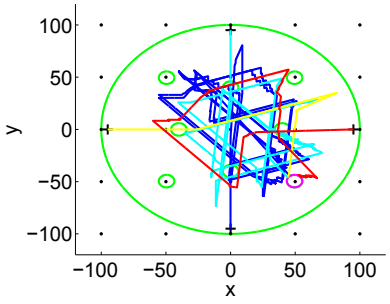


(a) Day 1

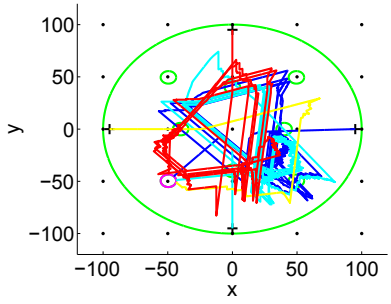
(b) Day 2



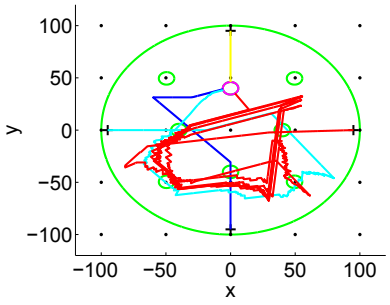
(c) Day 3



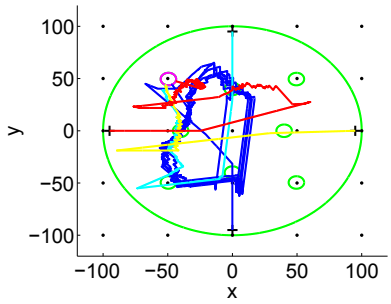
(d) Day 4



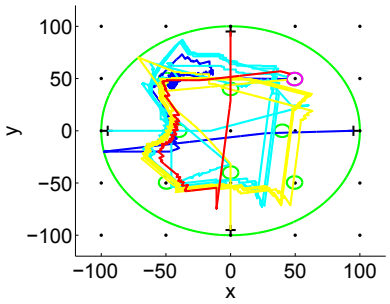
(e) Day 5



(f) Day 6

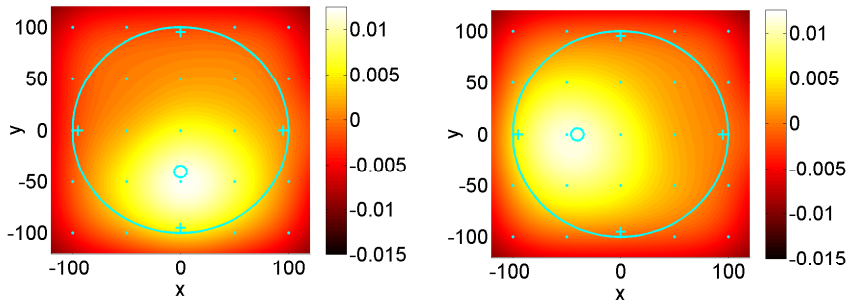


(g) Day 7



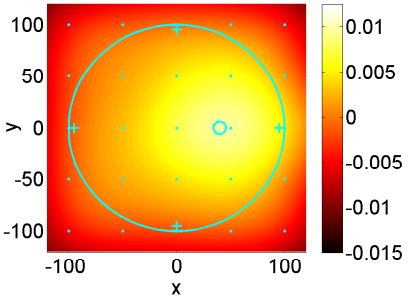
(h) Day 8

Figure 5.23. Colour-coded trajectories for 4 trials, from blue (trial 1) to red (trial 4), on the DMP task with increase in size and spacing of place cell firing fields, for each day. Black dots: centres of place cells firing fields, black crosses: starting locations. Magenta circle represents the chosen escape platform whereas green small circles are the other platforms. Simulation details and parameters are presented in Appendix J. Except here $N_P = 25$ and $\sigma_P = 50$.

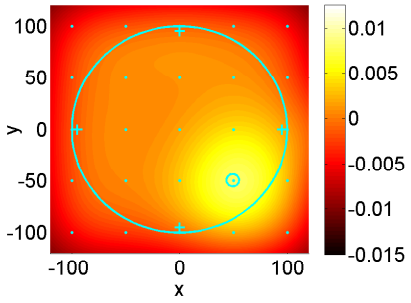


(a) Day 1

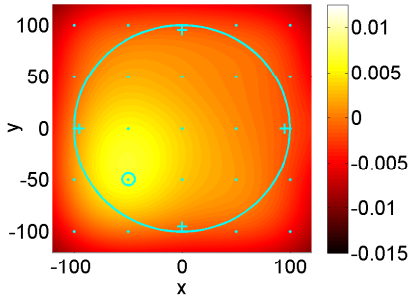
(b) Day 2



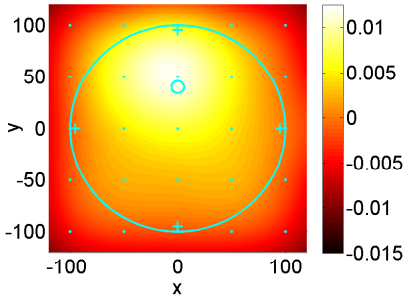
(c) Day 3



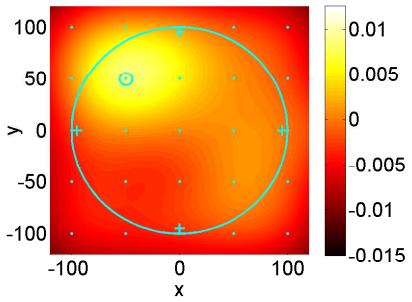
(d) Day 4



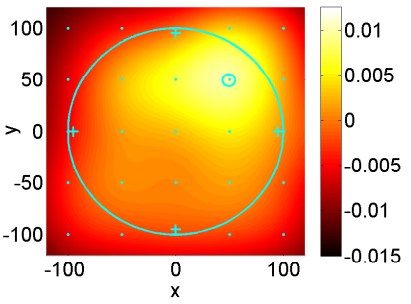
(e) Day 5



(f) Day 6

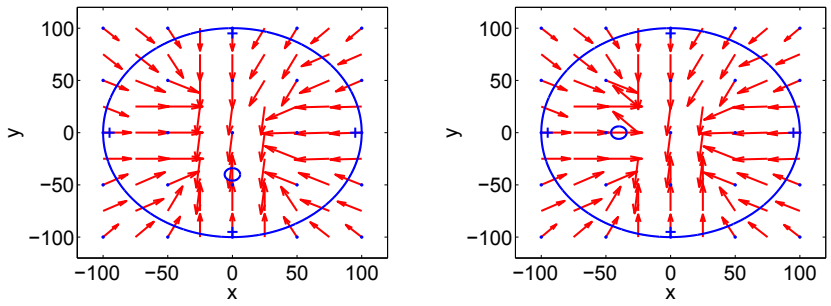


(g) Day 7



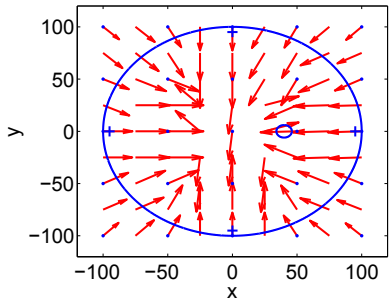
(h) Day 8

Figure 5.24. Value map of the DMP task with increase in size and spacing of place cell firing fields in Figure 5.23 after 4 trials for each day. Simulation details and parameters are presented in Appendix J. Except here $N_P = 25$ and $\sigma_P = 50$.

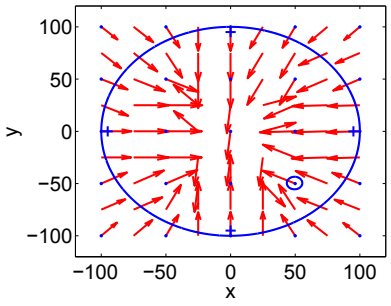


(a) Day 1

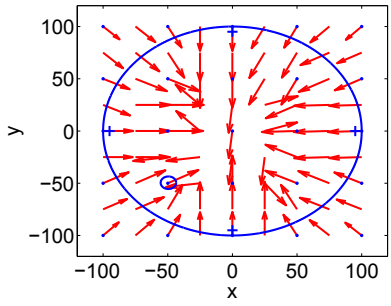
(b) Day 2



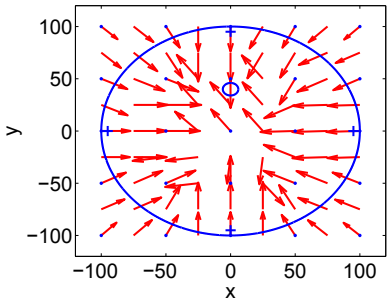
(c) Day 3



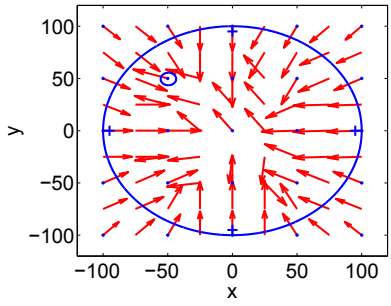
(d) Day 4



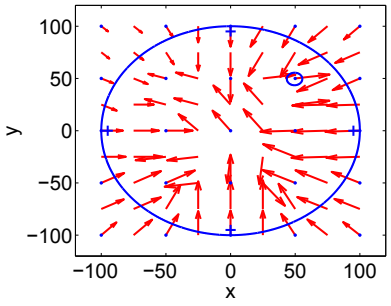
(e) Day 5



(f) Day 6

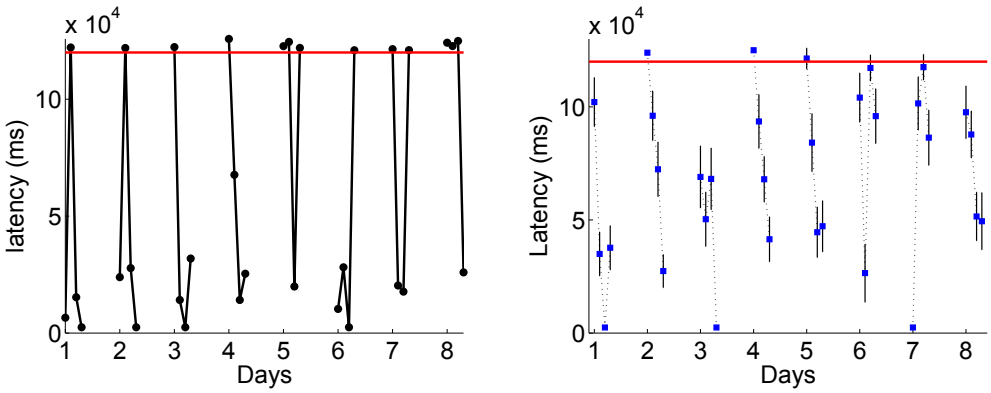


(g) Day 7



(h) Day 8

Figure 5.25. Policy map of the DMP task with increase in size and spacing of place cell firing fields in Figure 5.23 after 4 trials for each day. Simulation details and parameters are presented in Appendix J. Except here $N_P = 25$ and $\sigma_P = 50$.



(a) (b)

Figure 5.26. Plots of latencies to reach the platform for the DMP task with increase in size and spacing of place cell firing fields. (a) The latency of the agent in Figure 5.23 for the 4 different trails each day (dots) across 8 days (lines). Red line represents $t = t_{max}$. (b) Latency mean (blue squares) ± SEM (black lines) for 4 trials across 8 days. Here 15 artificial rats were simulated using the protocol in Appendix K.

5.5 DISCUSSION

In this chapter an actor-critic framework using firing rate neural models was proposed in order to simulate a DMP task. First, §5.3 showed that the rate version of the actor and critic neural networks, using elegibility traces as proposed by Doya [85], can achieve similar results as the model with spiking activity in [88] (first on a linear track and then on a square maze). Then, the model was tested on the DMP task in §5.4, for which an improvement of the latency time after the first trial in some cases was obtained. It was observed that the critic was able to learn a new goal location each day, but the performance of the actor was poor.

In order to improve the model performance, and based on experimental observations, visual cues were added in §5.4.2. This addition improved the learning by the actor. However, the implementation of visual cues did not improve the latency time on average (in comparison with the simulation without visual cues). A change in the distribution of the place cell firing field centres, size and

spacing, was implemented in §5.4.3. A decrease on the latency was observed, although the agent performance was worse at the first trial for some days.

5.5.1 *Further Work*

In order to improve the model, a sensitivity analysis of the model parameters should be undertaken. For example, the model presented in this chapter has lateral connectivity between actor neurons. Such connectivity allows reinforcement of actor neurons coding for similar directions of motion. Nevertheless, other models for the DMP task use independent actor neurons, e.g. Foster *et al.* [79]. Therefore, it would of interest to quantify the effect of the lateral connectivity strength of the model by varying ε_1 in (5.33) and analysing the resulting policy maps and latency time. Similarly, the performance of the model could be studied under variation of actor population size and range of encoded directions. Other measure of performance could be implemented, namely, the performance of the agent on a probe trial, implemented on animal tasks in [90]. On a probe trial, the platform does not come up until 60 s after starting the second trial of a given day. There are no modifications to the other trials in that day. In the probe trial the performance is measured as the time that the rat spends within a radius of the goal location. If the rat spends most of the time in this area, then it demonstrates a good search preference, in the sense that the animal does not reach the goal location randomly.

Furthermore, if $\varepsilon_1 = \sigma_n = 0$ in (5.33) then, following Doya [85], the policy can be implemented using (5.27) with

$$A(s(t); w^A) = \sum_{j=1}^{N_P} w_{lj}^{PA}(t) \frac{\rho_j^P(t)}{\sum_{i=1}^{N_P} \rho_i^P(t)}, \tag{5.39}$$

and the strength between place cell and actor neurons can be updated according to (5.28). Other methods of action selection, as the stochastic action selection

in [79], could be implemented. In such methods the swimming direction θ_l is chosen stochastically with probability P_l related to the actor activity:

$$P_l = \frac{\exp\left(2u_l^A(t)\right)}{\sum_{i=1}^{N_A} \exp\left(2u_i^A(t)\right)}. \tag{5.40}$$

The above method corresponds to the softmax action selection rules in [83].

The implementation of visual cues could be improved to better reflect experimental reality. Discussion with T. Bast, associate professor at the School of Psychology (University of Nottingham), indicates that after the first trial, the rats rapidly arrive at a zone near the goal location. In most experiments, rats stay in this zone until they reach the goal location. This can be implemented by making the agent rely on information about the selected cues at the beginning of trials 2 to 4 until it reaches the zone near the goal location, where the information about the environment is only obtained by place cell activity.

The action selection could be improved by assuming two components of responding to a stimulus, namely, the Pavlovian and the instrumental, as proposed in Huys *et al.* [92]. Here the stimulus can be thought of as the activity of place cells, and the response as the selection of the agent swimming direction. The difference between the two elements of responding is that the Pavlovian response is based on a learned expectation of gain or loss given a particular stimulus and is independent of the of the reward or punishment given to the agent as a consequence of its actions; whereas the instrumental response varies in time proportionally to the action outcome. Huys *et al.* proposed a variation of the learning rules in order consider the effects of Pavlovian and instrumental responses in learning. Such variation could be implemented in the model proposed in this chapter.

The distribution and size of the place cells firing fields could be improved in view of experimental data that suggests that the environment is represented by a finite continumm of scales of place cell firing fields, depending on the location of the cell in the hippocampus [38]. Moreover, as suggested in [83], learning methods for the RBFs can be applied, and such methods, as proposed by Poggio and Girosi in [93], change the centres and widths of place cells firing fields.

If these improvements translate into one trial learning of the DMP task, hypotheses about brain malfunction could be tested in the model by including brain damage, as in experiments done in [90]. Brain damage could be simulated by changing the learning parameters or adding noise to the place cell activity.

6

CONCLUSIONS AND FURTHER WORK

In this thesis mathematical and computational approaches have been used to understand brain function. Particularly, how intrinsic neural mechanisms can affect neural activity, as well as how neural activity can give rise to higher brain functions, such as navigation and spatial memory. It has been shown how mathematical tools and numerical simulations complement each other to gain further knowledge of a certain phenomena. For example, in Chapter 3 various models of grid cell activity were proposed in order to find a biophysically reasonable mechanism for long wavelength patterns. The linear stability of the proposed models was analysed and numerical continuation of the parameters was implemented in order to establish a relationship between the parameters and the wavelength of the resulting patterns. Importantly, the addition of dendritic processing and the implementation of axo-dendritic connections to the model resulted in a mechanism for long wavelength patterns. Results were tested against numerical simulations and showed excellent agreement. However, the proposed mechanism lacked experimental evidence.

Consequently, in Chapter 4 a IF network model of grid cell activity was proposed, based on experimental evidence. Firstly, this model was simulated in order to ascertain if it could exhibit the required activity pattern. Subsequently, the emergent travelling waves were described mathematically. Moreover, the Evans function, which determines wave stability, was derived and numerical continuation of the model parameters was performed. Importantly, it was shown that the maximum stable period of the travelling waves, that is believed to be

related to the scale of grid cell firing fields, can be controlled by the properties of the I_h current.

In Chapter 5 a navigation task using a reinforcement learning algorithm in a continuous actor-critic framework with firing rate models was presented. Although the learning rules for this algorithm were derived mathematically, the objective of this chapter was to describe how the brain is able to achieve navigation using computational operations in the spirit of Marr [10]. The agent was not able to show one trial learning, however, it was capable of learning new goal locations each day. Further components need to be added to the presented algorithm in order to achieve one trial learning.

Overall, the models proposed in this work described neural activity related to navigation and spatial memory at different levels. This neural activity was analytically studied and/or numerically simulated. Some models fail to provide the expected mechanisms, however their study allowed deeper understanding of the relationship between the model components and the resulting activity. Although in each chapter possible further work was outlined, in the following sections other generalisations are discussed.

6.1 EXTENSION TO 2D

The models presented in §3.7 and Chapter 4 were assumed to have a 1D somatic coordinate for numerical convenience. The emergence of travelling waves was observed in these models. Thus, they could be extended to 2D in order to investigate under which conditions travelling waves can arise. Moreover, the emergent patterns could be studied by applying a weakly nonlinear analysis (as in [8]) and therefore, the conditions under which hexagonal patterns are obtained, resembling those of the CAN models of grid cells, could be determined.

Actually, numerical simulations of the IF network model for grid cell dynamics in Chapter 4 have been extended to 2D in [69]. Due to the computational cost, these simulations were performed using GPUs. Under certain initial conditions, they show a periodic travelling wave moving downward. In order to study

such travelling wave solutions, first a change from spatio-temporal coordinates (x, y, t) to a travelling wave framework ξ must be performed. For example, in the work by Terman [94] this transformation is given by $\xi = (x, y) \cdot \mathbf{v} + \theta t$, where $\mathbf{v} \in \mathbb{R}^2$ is some unit vector, and $\theta > 0$ is the wave speed. Since the travelling wave observed in [69] moves downward, it can be assumed that $\mathbf{v} = (0, 1)$. This change of coordinates can be used to construct the travelling wave orbits and derive their stability.

6.2 SENSORY INFORMATION

One natural extension of this work is the integration of sensory information to the presented models. In Chapters 3 and 4 the proposed models can be driven by sensory input, whereas in Chapter 5 sensory input is needed by the agent to gain information about the environment and be able to perform a maze task.

Recently, theoretical approaches of how information integration could occur in the brain have been proposed [95, 96]. Most of these approaches are based on the head direction cell system but they can be extended to other systems. Generally, it is assumed that sensory inputs are ambiguous and that the brain is able to extract reliable information from these inputs.

One problem of integrating sensory information is selecting which of the inputs present at some time are the most reliable in order to achieve an optimal information integration. Jeffery *et al.* [95] studied how sensory information from different visual stimulus can be optimised to infer heading direction. In their model each sensory input (visual cue) is represented by a Gaussian curve, with centre at the heading direction activated by the sensory input and the height of the curve's peak denoting the intensity of such input. There are situations where the system receives input from two visual cues and the information of these cues has to be integrated to have optimal information about heading direction. These cues can be at different fixed locations or can change position between trials, thereby losing reliability. Jeffery *et al.* propose that because some sensory inputs are more reliable than others, the optimal information integration can

be achieved by strengthening the connections between the sensory neurons and head direction for a stable cue, and weakened otherwise. Here, a cue is stable if a set of sensory cells always activate the same set of head direction cells.

A Bayesian approach has been used by Zhang *et al.* [96] to infer heading direction using ambiguous information from two different sensory stimulus (related to different parts of the brain), optical flow and body acceleration. Their model uses two reciprocally connected CAN, each receiving information from different sensory stimulus. Each neuron in these networks codes for a preferred head direction value, uniformly distributed in parameter space. Zhang *et al.* show that optimal information integration is achieved by combined sensory inputs from different sources; moreover, the results obtained are improved when both sensory stimuli are presented. This type of approach can be used to infer the location on a maze by extending the description of the CAN to 2D.

A

CABLE EQUATION SOLUTION

In order to find (2.40), the Green's function for the cable equation (2.38) is solved with $I(\xi,t) = \delta(\xi)\delta(t)$, an impulse at $\xi = 0$ and $t = 0$. In this case (2.38) can be rewritten as

$$\frac{\partial V(\xi,t)}{\partial t} + \frac{V(\xi,t)}{\tau} - D\frac{\partial^2 V(\xi,t)}{\partial \xi^2} = \frac{r}{\tau}\delta(\xi)\delta(t). \tag{A.1}$$

Applying the Fourier transform with respect to the spatial variables gives

$$\frac{\partial \widetilde{V}(k,t)}{\partial t} + \frac{\widetilde{V}(k,t)}{\tau} + Dk^2\widetilde{V}(k,t) = \frac{r}{\tau}\delta(t), \tag{A.2}$$

where the *Fourier transform* and its inverse are defined by

$$\widetilde{f}(k) = \int_{-\infty}^{\infty} \mathrm{e}^{-ikx} f(x)\mathrm{d}x, \qquad f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{e}^{ikx}\widetilde{f}(k)\mathrm{d}k. \tag{A.3}$$

Note that (A.2) can be solved using an integrating factor. Thus

$$\widetilde{V}(k,t) = \frac{r}{\tau}\mathrm{e}^{-t\left(1/\tau + Dk^2\right)}H(t), \tag{A.4}$$

where $H(t)$ is the Heaviside function defined by (2.23). Now, applying the inverse Fourier transform to (A.4) gives

$$V(\xi,t) \;=\; \frac{r}{2\pi\tau}\mathrm{e}^{-t/\tau}H(t)\int_{-\infty}^{\infty}\mathrm{e}^{-tDk^2}\mathrm{e}^{ik\xi}\mathrm{d}k = \frac{r}{\tau}G_\infty(\xi,t), \tag{A.5}$$

with $G_\infty(\xi,t)$ as in (2.40).

B

AMPLITUDE EQUATIONS

In order to be able to predict the type of emergent patterns in the model proposed in §3.2, the relevant *amplitude equations* are derived following Chapter 1 in [8].

First, the system given by (3.4) with f the threshold-linear function, defined in (3.2), is analysed. Note that the linear stability analysis is the same as that in §3.2.1 because the quadratic terms are discarded, giving the dispersion relation (3.11) with $\gamma = f'(\overline{s}\overline{W} + I) = 1$. Since γ is fixed, g is chosen to be the bifurcation parameter, therefore $\widetilde{W}(k_c) = 1/g_c$. In order to derive the amplitude equations, a perturbation expansion for s is adopted

$$s = \overline{s} + \epsilon s_1 + \epsilon^2 s_2 + \epsilon^3 s_3 + \dots . \tag{B.1}$$

Time t is rescaled according to $T = \epsilon^2 t$, and g is set to $g = g_c + \epsilon^2 \Delta$. Here $|\epsilon| \ll 1$ and Δ is a measure of the distance from the bifurcation point. After equating powers of ϵ, the following hierarchy of equations is obtained

$$\overline{s} = g\left(\overline{s}\overline{W} + I\right), \tag{B.2a}$$

$$\mathcal{L}s_1 = 0, \tag{B.2b}$$

$$\mathcal{L}s_2 = \Delta\left(\overline{s}\overline{W} + I\right), \tag{B.2c}$$

$$\mathcal{L}s_3 = \Delta W \otimes s_1 - \tau\frac{\mathrm{d}s_1}{\mathrm{d}T}, \tag{B.2d}$$

where

$$\mathcal{L}s = s - g_c W \otimes s, \tag{B.3}$$

and $[W \otimes s](x,t) = \int_{\mathbb{R}^2} W\left(\mathbf{x}-\mathbf{x}'\right) s(\mathbf{x}',t)\mathrm{d}\mathbf{x}'$ denotes spatial convolution. Note that (B.2a) fixes the steady state and (B.2b) is linear, with solutions of the form

$$s_1(x,y,T) = A_1(T)\mathrm{e}^{ik_c x} + A_2(T)\mathrm{e}^{ik_c y} + c.c., \qquad A_j \in \mathbb{C}. \tag{B.4}$$

These solutions will form stripes if $A_1 = 0$ and $A_2 \neq 0$ (or vice versa) and spots if A_1, A_2 are both non-zero. The null-space of $\mathcal{L}$ is spanned by $\left\{\mathrm{e}^{\pm ik_c x}, \mathrm{e}^{\pm ik_c y}\right\}$.

Note that $\mathcal{L}s_n = h_n(\bar{s}, s_1, s_2, .., s_{n-1})$, therefore the Fredholm alternative method can be used to obtain solvability conditions from which the amplitudes A_1, A_2 can be found. It has been shown in [8] that $\mathcal{L}$ is self-adjoint with respect to the inner product defined by

$$\langle U, V\rangle = \frac{1}{|\Omega|}\int_\Omega U^*(\mathbf{r})V(\mathbf{r})\mathrm{d}\mathbf{r}, \tag{B.5}$$

with $\Omega = (0, 2\pi/k_c) \times (0, 2\pi/k_c)$ and * denotes complex conjugation. Note that

$$\langle s_1, \mathcal{L}s_n\rangle = \langle \mathcal{L}s_1, s_n\rangle = 0. \tag{B.6}$$

Therefore the solvability conditions are given by

$$\langle\left\{\mathrm{e}^{\pm ik_c x}, \mathrm{e}^{\pm ik_c y}\right\}, h_n\rangle = 0, \qquad n \geq 3. \tag{B.7}$$

Using the fact that $\langle \mathrm{e}^{ink_c x}, \mathrm{e}^{imk_c x}\rangle = \delta_{n,m}$ and $W \otimes s_1 = \widetilde{W}(k_c)s_1$, the following amplitude equations are obtained for the threshold linear firing rate function (for which $f''(\bar{s}) = 0$ and $f'''(\bar{s}) = 0$):

$$\tau\frac{\mathrm{d}A_1}{\mathrm{d}T} = \frac{\Delta}{g_c}A_1, \qquad \tau\frac{\mathrm{d}A_2}{\mathrm{d}T} = \frac{\Delta}{g_c}A_2. \tag{B.8}$$

In order to obtain the phase θ and the amplitude R of A, the system coordinates are changed to polar coordinates by taking $A_j = R_j\mathrm{e}^{i\theta_j}$, thus

$$\frac{\mathrm{d}R_1}{\mathrm{d}T} = \frac{\Delta}{\tau g_c}R_1, \qquad \frac{\mathrm{d}\theta_1}{\mathrm{d}T} = 0, \qquad \frac{\mathrm{d}R_2}{\mathrm{d}T} = \frac{\Delta}{\tau g_c}R_2, \qquad \frac{\mathrm{d}\theta_2}{\mathrm{d}T} = 0, \tag{B.9}$$

and the steady state is given by $(R_1, \theta_1, R_2, \theta_2) = (0, \theta_{1_0}, 0, \theta_{2_0})$ which implies that $A_1 = A_2 = 0$, suggesting that no stable patterns will form. Therefore, the theory fails to predict the spatial patterns formed by the model (3.4) with the threshold-linear firing rate. In fact, the theory above assumes that f is C^3.

Now, in order to predict the patterns generated by a simple model of grid cell activity in §3.2, f is changed to be a smooth differentiable the sigmoid function, described in (3.6), with $\gamma = f'\left(\overline{s}\overline{W} + I\right)$ as the bifurcation parameter, such that $\widetilde{W}(k_c) = 1/(g\gamma_c)$ and setting $\gamma = \gamma_c + \epsilon^2\Delta$, results in the following hierarchy equations

$$\overline{s} = gf\left(\overline{s}\overline{W} + I\right), \tag{B.10a}$$

$$\mathcal{L}s_1 = 0, \tag{B.10b}$$

$$\mathcal{L}s_2 = g\gamma_2 \left(W \otimes s_1\right)^2, \tag{B.10c}$$

$$\mathcal{L}s_3 = g\Delta W \otimes s_1 + 2g\gamma_2 \left(W \otimes s_1\right)\left(W \otimes s_2\right) + g\gamma_3 \left(W \otimes s_1\right)^3 - \tau\frac{\mathrm{d}s_1}{\mathrm{d}T}, \tag{B.10d}$$

with $\mathcal{L}$ as in (B.3) and $\gamma_2 = f''\left(\overline{s}\overline{W} + I\right)/2$, $\gamma_3 = f'''\left(\overline{s}\overline{W} + I\right)/6$. As above, the amplitude equations for A_1, A_2 are obtained using the solvability conditions. Note that $s_2 \sim s_1^2$, thus it can be assumed that

$$\begin{aligned} s_2 &= \alpha_0 + \alpha_1 \mathrm{e}^{2ik_c x} + \alpha_2 \mathrm{e}^{-2ik_c x} + \alpha_3 \mathrm{e}^{2ik_c y} + \alpha_4 \mathrm{e}^{-2ik_c y} + \alpha_5 \mathrm{e}^{ik_c(x+y)} \\ &+ \alpha_6 \mathrm{e}^{-ik_c(x+y)} + \alpha_7 \mathrm{e}^{ik_c(x-y)} + \alpha_8 \mathrm{e}^{ik_c(-x+y)} + \phi s_1. \end{aligned} \tag{B.11}$$

Therefore

$$\begin{aligned} W \otimes s_2 = \alpha_0\overline{W} + \widetilde{W}\left(2k_c\right)\left(\alpha_1 \mathrm{e}^{2ik_c x} + \alpha_2 \mathrm{e}^{-2ik_c x} + \alpha_3 \mathrm{e}^{2ik_c y} + \alpha_4 \mathrm{e}^{-2ik_c y}\right) \\ + \widetilde{W}\left(\sqrt{2}k_c\right)\left(\alpha_5 \mathrm{e}^{ik_c(x+y)} + \alpha_6 \mathrm{e}^{-ik_c(x+y)} + \alpha_7 \mathrm{e}^{ik_c(x-y)} + \alpha_8 \mathrm{e}^{ik_c(-x+y)}\right) \\ + \phi\widetilde{W}\left(k_c\right) s_1. \end{aligned} \tag{B.12}$$

The constant values of $\alpha_0, \ldots, \alpha_8$ are obtained by substituting (B.11) and (B.12) in (B.10c), and balancing terms to give

$$\begin{aligned} \alpha_0 &= \frac{2\gamma_2\left(|A_1|^2 + |A_2|^2\right)}{g\gamma_c^2\left(1 - g\gamma_c\overline{W}\right)}, & \alpha_1 &= \frac{\gamma_2 A_1^2}{g\gamma_c^2\left(1 - g\gamma_c\widetilde{W}\left(2k_c\right)\right)}, \\ \alpha_3 &= \frac{\gamma_2 A_2^2}{g\gamma_c^2\left(1 - g\gamma_c\widetilde{W}\left(2k_c\right)\right)} & \alpha_5 &= \frac{2\gamma_2 A_1 A_2}{g\gamma_c^2\left(1 - g\gamma_c\widetilde{W}\left(\sqrt{2}k_c\right)\right)}, \\ \alpha_7 &= \frac{2\gamma_2 A_1 A_2^*}{g\gamma_c^2\left(1 - g\gamma_c\widetilde{W}\left(\sqrt{2}k_c\right)\right)} & \alpha_8 &= \frac{2\gamma_2 A_1^* A_2}{g\gamma_c^2\left(1 - g\gamma_c\widetilde{W}\left(\sqrt{2}k_c\right)\right)}. \end{aligned} \tag{B.13}$$

The amplitude equations are then given by

$$\gamma_c\tau\frac{\mathrm{d}A_1}{\mathrm{d}T} = A_1\left(\Delta - \Upsilon\left|A_1\right|^2 - \Phi\left|A_2\right|^2\right), \tag{B.14a}$$

$$\gamma_c\tau\frac{\mathrm{d}A_2}{\mathrm{d}T} = A_2\left(\Delta - \Upsilon\left|A_2\right|^2 - \Phi\left|A_1\right|^2\right); \tag{B.14b}$$

where

$$\Upsilon = -\frac{1}{g\gamma_c^2}\left[2\gamma_2^2\left(\frac{2\overline{W}}{1-g\gamma_c\overline{W}} + \frac{\widetilde{W}(2k_c)}{1-g\gamma_c\widetilde{W}(2k_c)}\right) + \frac{3\gamma_3}{g}\right], \tag{B.15}$$

$$\Phi = -\frac{2}{g\gamma_c^2}\left[2\gamma_2^2\left(\frac{\overline{W}}{1-g\gamma_c\overline{W}} + \frac{2\widetilde{W}(\sqrt{2}k_c)}{1-g\gamma_c\widetilde{W}(\sqrt{2}k_c)}\right) + \frac{3\gamma_3}{g}\right]. \tag{B.16}$$

As above A_j is set to $A_j = R_j\mathrm{e}^{i\theta_j}$ to yield

$$\begin{aligned}\frac{\mathrm{d}R_1}{\mathrm{d}T} &= \frac{R_1}{\gamma_c\tau}\left(\Delta - \Upsilon R_1^2 - \Phi R_2^2\right), & \frac{\mathrm{d}\theta_1}{\mathrm{d}T} &= 0,\\ \frac{\mathrm{d}R_2}{\mathrm{d}T} &= \frac{R_2}{\gamma_c\tau}\left(\Delta - \Upsilon R_2^2 - \Phi R_1^2\right), & \frac{\mathrm{d}\theta_2}{\mathrm{d}T} &= 0.\end{aligned} \tag{B.17}$$

There are 3 solutions for this system: $R_1 = 0$, $R_2^2 = \Delta/\Upsilon$ with $\Upsilon > 0$ or vice versa, and $R_1^2 = R_2^2 = \Delta/\left(\Upsilon + \Phi\right)$ with $\Upsilon + \Phi > 0$. In order to predict which type of pattern will be formed in the system (3.4), conditions for the stability of (B.17) are derived. After linearising about the steady state and looking for solutions of the form $\chi(T) = \mathrm{e}^{\lambda T}\chi_0$, where χ is the perturbation vector and χ_0 is constant, the characteristic polynomial $\lambda^2\Lambda(\lambda) = 0$ is obtained, where

$$\begin{aligned}\Lambda(\lambda) &= \left(\tau\gamma_c\right)^2\lambda^2 - \tau\gamma_c\lambda\left(2\Delta - \left(3\Upsilon + \Phi\right)\left(R_1^2 + R_2^2\right)\right)\\ &+ \left(\Delta - 3\Upsilon R_1^2 - \Phi R_2^2\right)\left(\Delta - 3\Upsilon R_2^2 - \Phi R_1^2\right) - \left(2\Phi R_1 R_2\right)^2.\end{aligned} \tag{B.18}$$

Therefore, the solution that forms stripes ($A_1 = 0, A_2 \neq 0 \Rightarrow R_1 = 0, R_2 \neq 0$ or vice versa) will be stable when $\Phi > \Upsilon > 0$. Spotted patterns will be formed when $\Upsilon > \Phi$ and $\Upsilon + \Phi > 0$. In Figure 3.10 it is shown that spots or stripes are selected depending on the firing rate parameters values. For Figure 3.10a $\Phi = 20.9096 > \Upsilon = 10.669$, whereas for Figure 3.10b $\Upsilon = 12.1 > \Phi = 11.4904$. Spots are also selected in Figure 3.12, where $\Upsilon = 1.5227 > \Phi = -0.3367$ for Figure 3.12a, and $\Upsilon = 0.3678 > \Phi = 0.0435$ for Figure 3.12b.

C

SIMULATION DETAILS FOR A SIMPLE MODEL OF GRID CELL ACTIVITY

Here the 2D simulation method of the system (3.4) is presented. Let s_i denote the activity of the cell at position $\mathbf{x}_i$ where $\mathbf{x}_i = (x_i, y_i)$, $x_i, y_i \in 1, 2, ..., N$ (and therefore $\Delta x = 1$), with $N = 2^7$. Then the discrete version of the continuous system in (3.4) is given by:

$$\frac{\mathrm{d}s_i}{\mathrm{d}t} = \frac{1}{\tau}\left(-s_i + gf\left(\sum_j W_{ij}s_j + I\right)\right), \tag{C.1}$$

where

$$W_{ij} = W_0 H\left(R - \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}\right). \tag{C.2}$$

The domain is assumed to be periodic, so that the topology of the network is a torus. The system was simulated using an Euler method with $\Delta t = 1$ms and initialised with random activity. Figure 3.6b shows the resulting activity s after 500ms for a mesh of size $N \times N$.

D

LAPLACE TRANSFORM OF THE GREEN'S FUNCTION

In order to obtain equation (3.51) the Laplace transform, defined by equation (3.37), is applied to the system (3.42). This gives

$$\lambda \widehat{V}(\xi,\lambda) = -\frac{\widehat{V}(\xi,\lambda)}{\tau} + D\frac{\partial^2 \widehat{V}(\xi,\lambda)}{\partial \xi^2} + \rho \widehat{I}_s(\xi,\lambda). \tag{D.1}$$

Rearranging terms results in

$$\left(\varphi^2(\lambda) - \frac{\partial^2}{\partial \xi^2}\right)\widehat{V}(\xi,\lambda) = \frac{\rho}{D}\widehat{I}_s(\xi,\lambda), \qquad \varphi^2(\lambda) = \frac{\lambda + 1/\tau}{D}. \tag{D.2}$$

Now, changing variables to $\chi = \varphi(\lambda)\xi$ gives

$$\left(1 - \frac{\partial^2}{\partial \chi^2}\right)\widehat{V}(\chi,\lambda) = \frac{\rho}{D\varphi^2(\lambda)}\widehat{I}_s(\chi,\lambda). \tag{D.3}$$

The Green's function is found by applying the Fourier transform to the homogeneous problem of (D.3) (as in Appendix A), to give

$$(1+k^2)\widetilde{\widehat{V}}(k,\lambda) = 1. \tag{D.4}$$

The inverse Fourier transform is given by

$$\widehat{V}(\chi,\lambda) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{\mathrm{e}^{ik\chi}}{(1+ik)(1-ik)}\mathrm{d}k = i\mathrm{Res}\left(\frac{\mathrm{e}^{ik\chi}}{(1+ik)(1-ik)}\right). \tag{D.5}$$

Using the residue theorem yields

$$\widehat{G}(\chi) = \begin{cases} \mathrm{e}^{-\chi}/2 & \text{if } \chi > 0, \\ \mathrm{e}^{\chi}/2 & \text{if } \chi < 0 \end{cases} = \frac{\mathrm{e}^{-|\chi|}}{2}, \tag{D.6}$$

and the general solution is given by

$$\widehat{V}(\chi,\lambda) = 2\int_0^{\infty} \mathrm{d}\chi' \widehat{G}(\chi-\chi')\frac{\rho}{D\varphi^2(\lambda)}\widehat{I}_s(\chi',\lambda). \tag{D.7}$$

Changing to the original variables yields

$$\widehat{V}(\xi,\lambda) = \rho \int_0^\infty \mathrm{d}\xi' \widehat{G}(\xi-\xi',\lambda)\widehat{I}_s(\xi',\lambda), \tag{D.8}$$

with $\widehat{G}(\xi,\lambda)$ defined as in (3.51).

E

SIMULATIONS OF THE RESONANT MODEL WITH AXO-DENDRITIC CONNECTIONS

Here the 1D numerical simulation of the system (3.66) is presented. In order to be able to solve the system on a finite domain, the voltage variable is written in matrix form $V \in \mathbb{R}^{N\times M}$, where N is the dimension of the soma coordinate vector x and M the dimension of the dendrite coordinate vector ξ. Here $x \in [-a, a]$ where a is adjusted so that the domain comprise around five times the wavelength; $\xi \in [0, d]$ with d proportional to $\sqrt{D\tau}$ (the space constant in (2.38)). Thus, $V_{i,j}$ refers to the voltage at the dendritic position ξ_j of the cell at position x_i.

Finite differences are used to approximate the diffusion term in (3.66). Hence

$$\frac{\partial^2 V}{\partial \xi^2} \simeq \frac{V_{i,j+1} - 2V_{i,j} + V_{i,j-1}}{(\Delta\xi)^2}, \tag{E.1}$$

where $\Delta\xi = \xi_{j+1} - \xi_j$ for all j, thus $\Delta\xi = d/(M-1)$. Aperiodic boundary conditions were implemented in the dendritic space.

The injected current can be rewritten using the differential operator of the Green's function $\eta(t)$, which gives

$$\left(1 + \frac{1}{\alpha}\frac{\mathrm{d}}{\mathrm{d}t}\right)^2 A = \int_{\mathbb{R}} \mathrm{d}x'\, W(x - x', \xi) f\left(h\left(x', t\right)\right), \quad A = I_{inj} - I_{ext}(\xi). \tag{E.2}$$

Therefore, the system given in (3.66) can be approximated by

$$\begin{aligned}
\frac{\mathrm{d}V_{i,j}}{\mathrm{d}t} &= -\frac{V_{i,j}}{\tau} + D\frac{V_{i,j+1} - 2V_{i,j} + V_{i,j-1}}{(\Delta\xi)^2} - \frac{1}{C}\left(I_{i,j} - gI_{inj_{i,j}}\right), \\
\frac{\mathrm{d}I_{i,j}}{\mathrm{d}t} &= \frac{1}{L}\left(-rI_{i,j} + V_{i,j}\right), \\
\frac{\mathrm{d}A_{i,j}}{\mathrm{d}t} &= \alpha\left(\widetilde{A}_{i,j} - A_{i,j}\right), \\
\frac{\mathrm{d}\widetilde{A}_{i,j}}{\mathrm{d}t} &= \alpha\left(B_{i,j} - \widetilde{A}_{i,j}\right),
\end{aligned}$$

where

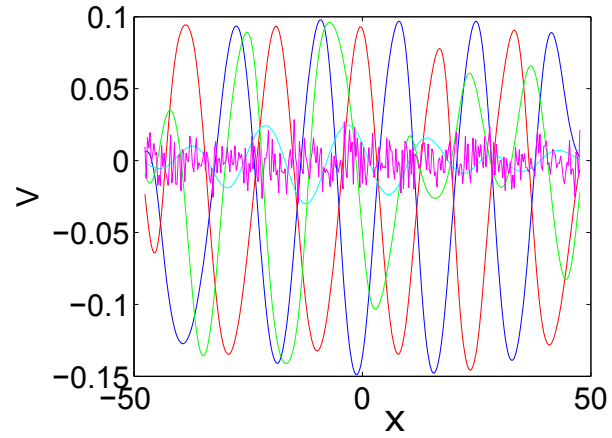
$$\left(1 + \frac{1}{\alpha}\frac{\mathrm{d}}{\mathrm{d}t}\right) A_{i,j} = \widetilde{A}_{i,j}, \qquad I_{inj_{i,j}} = A_{i,j} + I_{ext_j}, \tag{E.3}$$

and

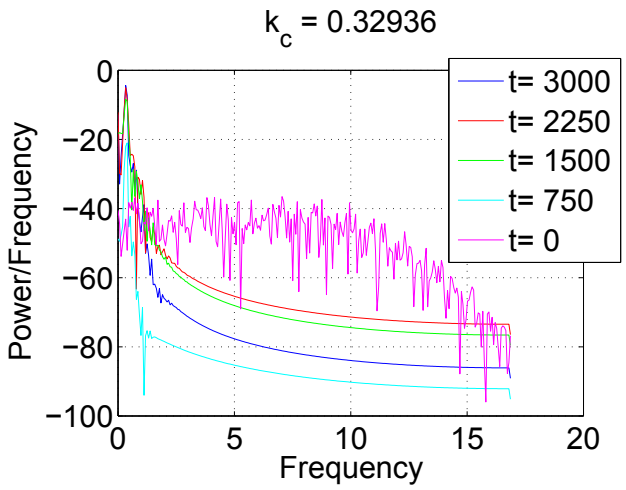
$$B_{i,j} = \sum_{l=1}^{N} W(x_i - x_l, \xi_j) f(h_l) \Delta x, \tag{E.4}$$

with $h_l = V(x_l, \xi = 0)$ and $\Delta x = 2a/N$. Note that $I, A, \widetilde{A}, B \in \mathbb{R}^{N\times M}$, whereas $f, g_{ext} \in \mathbb{R}^N$ are measured just at the soma $V(x_i = 0, \xi_j)$. Aperiodic boundary conditions were implemented in the somatic space.

MATLAB's `ode45` solver was used to evolve the system from the initial conditions $(V_{i,j}, I_{i,j}, A_{i,j}, \widetilde{A}_{i,j}) = (0, 0, -I_{ext_j}, -I_{ext_j})$. Figure 3.26 shows the evolution of the voltage at the soma. Figure E.1a shows the voltage at the soma for different times whereas Figure E.1b gives the wave-number calculated using the power spectrum (see caption for details). For the chosen parameters the resulting value of k_c is $k_c = 0.33$ (see Fig. 3.27), showing that the simulation and the theory agree. Moreover, it demonstrates that the numerical choices (domain size for the somatic and dendritic coordinate, Δx and $\Delta\xi$) are right.



(a)



(b)

Figure E.1. (a): Plot of the voltage at the soma ($\xi = 0$) taken at different times (colour-coded) from the simulation in Figure 3.26. (b): Periodogram using Fourier Transforms on the voltage at the soma, at different times (as in (b)). The periodogram shows a peak at the preferred spatial frequency. Around $t = 750$ there is a clear frequency preference at $k_c = 0.32936$ in the simulation.

F

SIMULATIONS OF AN INTEGRATE-AND-FIRE NETWORK MODEL FOR GRID CELL DYNAMICS

In order to solve (4.8) on a finite domain in $x \in (-L, L)$, x is written as a vector in $\mathbb{R}^N$, with $V, n_h, \psi \in \mathbb{R}^N$. As in Appendix E, the differential operator for η is used, and the discretization of ψ is evolved according to

$$\left(1 + \frac{1}{\alpha}^2 \frac{\mathrm{d}}{\mathrm{d}t}\right) \psi_i = \sum_j W_{ij} \sum_{m \in \mathbb{Z}} \delta(t - T^m(x_j)) \Delta x, \tag{F.1}$$

where x_j is the position of the jth cell, $\Delta x = x_j - x_{j-1}, \forall j$, with $\Delta x = 2L/(N-1)$. As in [97] it is convenient to write ψ_i as the solution to

$$\frac{1}{\alpha} \frac{\mathrm{d}\psi_i}{\mathrm{d}t} = y_i - \psi_i, \tag{F.2}$$

$$\frac{1}{\alpha} \frac{\mathrm{d}y_i}{\mathrm{d}t} = -y_i, \tag{F.3}$$

with y_i discontinuously updated according to $y_i \to y_i + \alpha W_{ij} \Delta x$ at times $T^m(x_j)$, and $W_{ij} = W(x_i - x_j)$. Table 1 shows the pseudocode that was implemented in MATLAB for the system $X = [V; n_h; \psi; y]$, where

$$\frac{\mathrm{d}V_i}{\mathrm{d}t} = \frac{1}{C} \left(-g_l V_i + g_h n_{h,i} + g_{syn} \psi_i\right), \tag{F.4}$$

$$\frac{\mathrm{d}n_{h,i}}{\mathrm{d}t} = \frac{1}{\tau_h} \left(n_{h,\infty}(V_i) - n_{h,i}\right), \tag{F.5}$$

with ψ_i given by (F.2) and y_i by (F.3). The evolution of the system X was done using the RK4 method.

Comments: $X(t)$ denotes the state of the network at time t.
$\mathcal{F}$ is the list of the firing events and their locations.
t^R is a vector with entries storing the amount of time that each neuron has been in the refractory state (0 otherwise).
t^s is given by $V(t^s) = V_{th}$.
Pseudocode:
Initialise network $t = 0$, $X(t) = 0$, $t^R = 0$.
while $t \leq t_{max}$ **do**
  Set $t_{new} = t + \Delta t$
  Set $\mathcal{L} = \emptyset$
  Evolve X from t to t_{new}
  for $j = 1, N$ **do**
    if $V_j(t_{new}) \geq V_{th}$ *and* $t_j^R = 0$ **then**
      Add $V_j(t_{new})$ to $\mathcal{L}$
    end
  end
  if $\mathcal{L} \neq \emptyset$ **then**
    Set $\mathcal{H} = \emptyset$
    Interpolate $V(t_{new}) \in \mathcal{L}$
    Find $t^* = \min_{\mathcal{L}} t^s$
    Set $t_{new} = t^*$
    Evolve X from t to t_{new}
    for $j = 1, N$ **do**
      if $V_j(t_{new}) \geq V_{th}$ *and* $t_j^R = 0$ **then**
        Add (t_j, x_j) to $\mathcal{F}$
        Add j to $\mathcal{H}$
        for $i = 1, N$ **do**
          Set $y_i \to y_i + \alpha \Delta x W_{ij}$
        end
      end
    end
  end
  Set $\delta = t_{new} - t$
  for $j = 1, N$ **do**
    if $j \in \mathcal{H}$ *or* $t_j^R \neq 0$ **then**
      Set $t_j^R = t_j^R + \delta$
      Set $V_j(t_{new}) = V_r$
      if $t_j^R \geq \tau_R$ **then**
        Set $t_j^R = 0$
      end
    end
  end
  Set $t = t_{new}$
end

Algorithm 1: Evolution of the network

G

DERIVATION OF THE TD ERROR AND LEARNING RULES

Here the TD error and learning rules for the maze task in §5.3 are derived. First, the solution for the critic evolution (5.30) is approximated by

$$u_i^C(t) = \int_0^t G^C(t-t')g_i(t')\mathrm{d}t' = (G^C * g_i)(t), \tag{G.1}$$

where $*$ represents a temporal convolution and G^C is the Green's function for the differential operator Q^C, given by

$$G^C(t) = \varepsilon_0 \frac{\mathrm{e}^{-t/\tau_m} - \mathrm{e}^{-t/\tau_s}}{\tau_m - \tau_s} H(t), \tag{G.2}$$

with $H(t)$ the Heaviside step function and $g_i(t)$ taken from (5.30),

$$g_i(t) = \sum_{j=1}^{N_P} w_{ij}^{PC}(t)\rho_j^P(t). \tag{G.3}$$

The solution for the critic activity allows the calculation of the derivative of the value function

$$\frac{\mathrm{d}V(\mathbf{x}(t))}{\mathrm{d}t} = \frac{\mathrm{d}\left(\frac{v}{N_C}\sum_{i=1}^{N_C}\rho_i^C(t) + V_0\right)}{\mathrm{d}t} = \frac{v}{N_C}\sum_{i=1}^{N_C}\frac{\mathrm{d}\rho_i^C(t)}{\mathrm{d}t}, \tag{G.4}$$

where

$$\frac{\mathrm{d}\rho_i^C(u_i^C(t))}{\mathrm{d}t} = \frac{\partial\rho_i^C}{\partial u_i^C}\frac{\partial u_i^C}{\partial t} = \frac{\partial\rho_i^C}{\partial u_i^C}\frac{\partial(G^C * g_i)(t)}{\partial t} = \frac{\partial\rho_i^C}{\partial u_i^C}\left(\frac{\partial G^C}{\partial t} * g_i\right)(t). \tag{G.5}$$

Now, defining

$$\psi_i(t) = \left(\frac{\partial G^C}{\partial t} * g_i\right)(t), \tag{G.6}$$

and taking the Laplace transform, as defined by (3.37), results in

$$\widehat{\psi}_i(\lambda) = (\lambda \widehat{G}^C(\lambda) - G^C(0))(\widehat{g}_i(\lambda)) = \lambda \widehat{G}^C(\lambda)\widehat{g}_i(\lambda). \tag{G.7}$$

The last term in the above equation was obtained by applying the the convolution theorem and assuming $G^C(0) = 0$. Here $\widehat{G}^C(\lambda)$ is given by

$$\widehat{G}^C(\lambda) = \frac{\varepsilon_0}{(1+\lambda\tau_m)(1+\lambda\tau_s)}. \tag{G.8}$$

Rearranging the terms in (G.7) and taking the inverse Laplace transform yields

$$\left(1+\tau_m\frac{\mathrm{d}}{\mathrm{d}t}\right)\left(1+\tau_s\frac{\mathrm{d}}{\mathrm{d}t}\right)\psi_i(t) = \varepsilon_0\frac{\mathrm{d}g_i}{\mathrm{d}t}. \tag{G.9}$$

Substitution of (5.32) and (G.4) in (5.20) gives the TD error

$$\begin{aligned}\delta(t) &= \frac{v}{N_C}\sum_{i=1}^{N_C}\frac{\partial\rho_i^C}{\partial u_i^C}\psi_i(t) - \frac{1}{\tau_r}\left(\frac{v}{N_C}\sum_{i=1}^{N_C}\rho_i^C(t) + V_0\right) + r(\mathbf{a}(t),\mathbf{x}(t)) \\ &= \frac{v}{N_C}\sum_{i=1}^{N_C}\left(\frac{\partial\rho_i^C}{\partial u_i^C}\psi_i(t) - \frac{1}{\tau_r}\rho_i^C(t)\right) - \frac{V_0}{\tau_r} + r(\mathbf{a}(t),\mathbf{x}(t)).\end{aligned} \tag{G.10}$$

The parameter update from (5.25) and (5.26), are now given by

$$\frac{\mathrm{d}w_{ij}^{PC}(t)}{\mathrm{d}t} = \eta_c\delta(t)e_{ij}(t), \tag{G.11}$$

$$\frac{\mathrm{d}e_{ij}(t)}{\mathrm{d}t} = \frac{1}{\kappa}e_{ij}(t) + \frac{\partial V(\mathbf{x}(t);w^{PC})}{\partial w_{ij}^{PC}}. \tag{G.12}$$

Note that the eligibility trace e has the same dimensions as the parameter matrix w^{PC}. Here the derivative of the value function approximation with respect to the parameters is given by

$$\frac{\partial V(\mathbf{x}(t))}{\partial w_{ij}^{PC}} = \frac{\partial}{\partial w_{ij}^{PC}}\left[\frac{v}{N_C}\sum_{k=1}^{N_C}\rho_k^C(t) + V_0\right] = \frac{v}{N_C}\sum_{k=1}^{N_C}\frac{\partial\rho_k^C(t)}{\partial w_{ij}^{PC}}, \tag{G.13}$$

where

$$\frac{\partial\rho_k^C(u_k^C(t))}{\partial w_{ij}^{PC}} = \frac{\partial\rho_k^C}{\partial u_k^C}\frac{\partial u_k^C}{\partial w_{ij}^{PC}}. \tag{G.14}$$

The derivative of the critic activity with respect to the weights is calculated using the solution for the critic activity (G.1) in the following way

$$\begin{aligned}\frac{\partial u_k^C}{\partial w_{ij}^{PC}} &= \int_0^t G^C(t-t')\sum_{l=1}^{N_P}\frac{\partial w_{kl}^{PC}(t')}{\partial w_{ij}^{PC}}\rho_l^P(t')\mathrm{d}t' \\ &= \int_0^t G^C(t-t')\sum_{l=1}^{N_P}\delta_{ki}\delta_{jl}\rho_l^P(t')\mathrm{d}t' \\ &= \delta_{ki}\int_0^t G^C(t-t')\rho_j^P(t')\mathrm{d}t' = \delta_{ki}\left(G^C * \rho_j^P\right)(t), \end{aligned} \tag{G.15}$$

where

$$\delta_{ij} = \begin{cases} 1, & \text{if } i = j \\ 0, & \text{if } i \neq j \end{cases}. \tag{G.16}$$

Substitution into (G.13) gives

$$\frac{\partial V(\mathbf{x}(t))}{\partial w_{ij}^{PC}} = \frac{v}{N_C}\sum_{k=1}^{N_C}\frac{\partial \rho_k^C}{\partial u_k^C}\delta_{ki}\left(G^C * \rho_j^P\right)(t) = \frac{v}{N_C}\frac{\partial \rho_i^C}{\partial u_i^C}\left(G^C * \rho_j^P\right)(t). \tag{G.17}$$

In order to be able to computationally implement the above equation, let $\xi_j(t) = \left(G^C * \rho_j^P\right)(t)$ and apply the Laplace transform to find

$$\widehat{\xi}_j(\lambda) = \left(\widehat{G}^C(\lambda)\right)\left(\widehat{\rho}_j^P(\lambda)\right). \tag{G.18}$$

After substituting $\widehat{G}^C(\lambda)$ with (G.8) and rearranging terms gives

$$(1+\lambda\tau_m)(1+\lambda\tau_s)\,\widehat{\xi}_j(\lambda) = \varepsilon_0\widehat{\rho}_j^P(\lambda), \tag{G.19}$$

which upon inverse Laplace transform yields:

$$\left(1+\tau_m\frac{\mathrm{d}}{\mathrm{d}t}\right)\left(1+\tau_s\frac{\mathrm{d}}{\mathrm{d}t}\right)\xi_j(t) = \varepsilon_0\rho_j^P(t). \tag{G.20}$$

H

LINEAR TRACK SIMULATION

In this simulation 20 trials were performed. The environment is a narrow rectangle of size 40 cm $\times$ 4 cm centred around the origin. A trial starts when the agent is put at a starting position and ends when the goal location is reached. The starting point is at $\mathbf{x} = (x, y) = (-17.5, 0)$ and the goal location is at $x \geq 16$. For each trial the agent obtains the activity of the place cell by applying (5.29) to its current position. Because the action choice is clamped to go directly towards the goal location placed to the right of the initial position, the state evolves as

$$\frac{\mathrm{d}\mathbf{x}(t)}{\mathrm{d}t} = s\begin{pmatrix}1\\0\end{pmatrix}, \tag{H.1}$$

where $s > 0$ is constant speed and therefore the action taken is given by $\mathbf{a} = (\cos(x), \sin(y))$. Then, the activity of the critic is updated using (5.30). For computational convenience a new variable $z^C \in \mathbb{R}^{N_C}$ is introduced which gives the following system:

$$\frac{\mathrm{d}u_i^C(t)}{\mathrm{d}t} = \frac{1}{\tau_s}\left(z_i^C(t) - u_i^C(t)\right), \tag{H.2}$$

$$\frac{\mathrm{d}z_i^C(t)}{\mathrm{d}t} = \frac{1}{\tau_m}\left(\varepsilon_0 g_i(t) - z_i^C(t)\right), \tag{H.3}$$

with $i = 1, 2, \ldots, N_C$. Here $g(t)$ is given by (G.3). After updating the critic activity, the TD error (G.10) is calculated, with the critic firing rate obtained by applying (5.31) to $u^C(t)$ and the activity of $\psi(t)$ is given by (G.9). Expanding and rearranging terms results in

$$\psi_i(t) + (\tau_m + \tau_s)\frac{\mathrm{d}\psi_i(t)}{\mathrm{d}t} + \frac{\mathrm{d}q_i(t)}{\mathrm{d}t} = 0, \tag{H.4}$$

where $q_i(t) = \tau_m\tau_s \dfrac{\mathrm{d}\psi_i(t)}{\mathrm{d}t} - \varepsilon_0 g_i(t)$. This gives the following system

$$\frac{\mathrm{d}\psi_i(t)}{\mathrm{d}t} = \frac{q_i(t) + \varepsilon_0 g_i(t)}{\tau_m\tau_s}, \tag{H.5}$$

$$\frac{\mathrm{d}q_i(t)}{\mathrm{d}t} = -\psi_i(t) - (\tau_m + \tau_s)\left(\frac{q_i(t) + \varepsilon_0 g_i(t)}{\tau_m\tau_s}\right). \tag{H.6}$$

The reward rate term $r(t)$ is given by the difference of two decaying traces:

$$r(t) = \frac{r_a(t) - r_b(t)}{\tau_a - \tau_b}, \tag{H.7}$$

where

$$\frac{\mathrm{d}r_a(t)}{\mathrm{d}t} = -\frac{r_a}{\tau_a}, \tag{H.8}$$

$$\frac{\mathrm{d}r_b(t)}{\mathrm{d}t} = -\frac{r_b}{\tau_b}. \tag{H.9}$$

The reward R instantaneously update the traces when the agent reaches the goal ($R > 0$) or when the agent bumps into a wall ($R < 0$) in the following way

$$r_a(t) \to r_a(t) + R, \qquad r_b(t) \to r_b(t) + R. \tag{H.10}$$

The critic weights' evolution is described by (G.11) and (G.12). In order to calculate the final term on the right-hand side of (G.12), (G.20) is used. Here, γ is introduced so that the system can be solved numerically. Thus

$$\frac{\mathrm{d}\xi_j(t)}{\mathrm{d}t} = \frac{1}{\tau_s}\left(\gamma_j(t) - \xi_j(t)\right), \tag{H.11}$$

$$\frac{\mathrm{d}\gamma_j(t)}{\mathrm{d}t} = \frac{1}{\tau_m}\left(\varepsilon_0\rho_j^P(t) - \gamma_j(t)\right), \tag{H.12}$$

with $j = 1, 2, \ldots, N_P$. Substitution into (G.12) gives

$$\frac{\mathrm{d}e_{ij}(t)}{\mathrm{d}t} = \frac{1}{\kappa}e_{ij}(t) + \frac{v}{N_C}\frac{\partial\rho_i^C}{\partial u_i^C}\xi_j(t). \tag{H.13}$$

Parameters:

- For the place cell firing rate function (5.29): $\rho_p = 1$ Hz, $\sigma_P = 2$ cm and the place cell centres are arranged on a grid with spacing σ_P that covers all the environment plus a distance of σ_P. In this case there are 115 centres, so that $N_P = 115$.

- For the critic dynamics (5.30): $\tau_m = 20$ ms, $\tau_s = 5$ ms, $\varepsilon_0 = 1$ mV×ms.
- For the critic firing rate (5.31): $\mu = 0.5$ Hz, $\beta = 0.1$ mV^{-1}, $\Theta = 0$ mV.
- For the value function (5.32): $v = 1$ [reward units], $N_C = 100$ and $V_0 = -0.515$ [reward units] $\times$ ms.
- For the TD error (5.20): $\tau_r = 3000$ ms. The rates for the reward traces are $\tau_a = 200$ ms for (H.8) and $\tau_b = 10$ ms for (H.9), whereas the instantaneous reward when reaching the goal location is $R = 200$ [reward units]. The time constant of the eligibility trace (5.26) is $\kappa = \tau_r$. For the critic learning rate value in the synaptic weight update (5.25) $\eta_C = 50$ ms[reward units]$^{-1}$mV^{-1} is chosen.

The system was evolved using Euler's method with a time-step $\Delta t = 0.2$ms. The system was solved simultaneously for all the variables: $\mathbf{x} \in \mathbb{R}^2$, $u^C, z^C, \psi, q \in \mathbb{R}^{N_C}$, $r_a, r_b \in \mathbb{R}$, $\xi, \gamma \in \mathbb{R}^{N_P}$, and the weight and eligibility matrices $w^{PC}, e \in \mathbb{R}^{N_C \times N_P}$. The speed of the agent is $s = 0.005$cm/ms. All the variables were initialised with zero activity, except the weight matrix whose entries were randomly drawn from a normal distribution with mean equal to 0.1 and standard deviation of 0.005. The weights were constrained to the range $0 \leq w^{PC} \leq 0.65$, to avoid negative or runaway values.

For the value map in Figure 5.7a the value function (5.32) was calculated for every point on a spatial mesh with spacing of 0.1 cm.

I

SIMULATION OF THE MAZE TASK

Here the maze is a square area of 20 cm $\times$ 20 cm centred around the origin. The goal is a circle located at the origin with radius of 1 cm surrounded by a U-shaped obstacle of size 10 cm $\times$ 2 cm on three sides. As in the linear track simulation, the place cell centres are arranged on a grid with spacing σ_P.

Here 20 trials are simulated. At each trial, the starting position was randomly selected from four possible positions at $\mathbf{x} \in \{(\pm 7.5, 0), (0, \pm 7.5)\}$ cm. The state evolution is given by (5.37) and the place cell activity is obtained by applying (5.29) to the agent's current position. The critic is simulated as in Appendix H. For the actor dynamics $z^A \in \mathbb{R}^{N_A}$ is introduced:

$$\frac{\mathrm{d}z_l^A(t)}{\mathrm{d}t} = \frac{1}{\nu_\vartheta}\left(z_l^A(t) - u_l^A(t)\right), \tag{I.1}$$

$$\frac{\mathrm{d}u_l^A(t)}{\mathrm{d}t} = \frac{1}{\tau_\vartheta}\left(\varepsilon_0 h_l - z_l^A(t)\right), \tag{I.2}$$

$$h_l = \sum_{j=1}^{N_P} w_{lj}^{PA}(t)\rho_j^P(t) + \frac{\varepsilon_1}{\varepsilon_0}\sum_{k=1}^{N_A} w_{lk}^{AA}(t)u_k^A(t) + \frac{\sigma_n}{\varepsilon_0}n_l(t), \tag{I.3}$$

where n evolves according to (5.35). The connectivity within the lth and kth action cells is given by

$$w_{lk} = \frac{w_-}{N_A} + w_+\frac{f(l,k)}{\sum_k f(l,k)}, \qquad f(l,k) = (1-\delta_{lk})\exp(\zeta\cos(\theta_l - \theta_k)), \tag{I.4}$$

where $\theta_l = 2l\pi/N_A$, for $l = 1, 2, \ldots, N_A$.

The synaptic weights between place and actor cells are updated via the method outlined for the critic updates in Appendix H, which gives

$$\frac{\mathrm{d}\xi_j^A(t)}{\mathrm{d}t} = \frac{1}{\nu_\vartheta}\left(\gamma_j^A(t) - \xi_j^A(t)\right), \tag{I.5}$$

$$\frac{\mathrm{d}\gamma_j^A(t)}{\mathrm{d}t} = \frac{1}{\tau_\vartheta}\left(\varepsilon_0\rho_j^P(t) - \gamma_j^A(t)\right), \tag{I.6}$$

with $j = 1, 2, \ldots, N_P$. Therefore

$$\frac{\mathrm{d}w_{lj}^{PA}(t)}{\mathrm{d}t} = \eta_A\delta(t)\frac{\partial\rho_l^A}{\partial u_l^A}\xi_j^A(t). \tag{I.7}$$

Here an instantaneous negative reward R_{obs} was given every time that the agent collides with the wall. On these occasions, the state was updated according to $\Delta x\delta(t)\Psi(\mathbf{x}(t))$, where $\Psi(\mathbf{x}(t))$ is a unit vector that points perpendicularly inward from the obstacle surface, and Δx a constant. A trial ends when the goal location is reached and an instantaneous reward R_{goal} is given, or when $t = t_{max} = 50$ ms. The simulations were run simultaneously using Euler's method ($\Delta t = 0.2$ ms) for $\mathbf{x} \in \mathbb{R}^2$, $u^C, z^C, \psi, q \in \mathbb{R}^{N_C}$, $r_a, r_b \in \mathbb{R}$, $\xi, \gamma, \xi^A, \gamma^A \in \mathbb{R}^{N_P}$, $u^A, z^A, n \in \mathbb{R}^{N_A}$, and the weight and eligibility matrices $w^{PC}, e \in \mathbb{R}^{N_C \times N_P}$, and $w^{PA} \in \mathbb{R}^{N_A \times N_P}$. As in the linear track case, all the variables were initiated with zero activity except the weight matrices whose entries were randomly drawn from a normal distribution with mean equal to 0.1 and standard deviation of 0.005. Here $0 \leq w^{PC}, w^{PA} \leq 0.65$.

Parameters for the actor are: $N_A = 180$, $\nu_\vartheta = 20$ ms, $\tau_\vartheta = 50$ ms and $\varepsilon_1 = 1/30$ mV $\times$ ms. For the lateral connections $w_+ = 30$, $w_- = -60$, $\zeta = 8$. The learning rate $\eta_A = 5$ ms[reward units]$^{-1}$mV^{-1}. For the noise component $\tau_n = 1$ ms, $\sigma_{n0} = 1$, $V_{max} = 0.1$ mV, and $V_{min} = -0.515$ mV. The speed of the agent is $s = 0.01$cm/ms and $\Delta x = 5$ cm. Due to the change of dimensions relative to the linear track problem N_P increases to 169. Here $\eta_C = 20$ ms[reward units]$^{-1}$mV^{-1}, and for the TD error $\tau_r = 500$ ms, $\kappa = 2\tau_r/3$. The rewards are $R_{goal} = 400$ [reward units] and $R_{obs} = -0.00005$ [reward units]. Other parameters are as in Appendix H.

The policy map was obtained by obtained by choosing the direction of the cell with maximum $h'_l = \rho^A(\sum_{j=1}^{N_P} w_{lj}^{PA}\rho_j^P)$ times the sum of h'. For plotting, MATLAB's `quiver` function was used.

J

SIMULATION OF THE DMP TASK

Here the DMP task is simulated by solving the same system as in Appendix I. The environment has the same dimensions as in [90]. The maze has a 1 m radius and the escape platform each have a 12 cm diameter, distributed along two concentric circles with radii of 70 cm and 40 cm. The centres of the escape platforms are located at $\mathbf{x} \in \{(\pm p_1, \pm p_1), (p_2, \pm p_3), (\pm p_3, p_2)\}$ where $p_1 = 49.4975$ cm, $p_2 = 0$ cm and $p_3 = 40$ cm. The place cell firing fields centres are located in concentric circles with σ_P separation between and within the circles. Here $N_P = 58$, $\sigma_P = 25$ cm. The starting positions are at $\mathbf{x} \in \{(\pm 95, 0), (0, \pm 95)\}$.

If the rat is not able to reach the goal location after $t_{max} = 120$ s, then it is guided towards it: a straight line between the goal location and the rat current position is traced and the following actions are forced to follow this line. The weights between the place and the actor cells are updated according to this forced action, so that (5.38) is given by

$$\dot{w}_{lj}^{PA} = \eta_A \delta(t) \Upsilon_l \frac{\partial \rho^A(u_l^A; w^{PA})}{\partial w_{lj}^{PA}}, \tag{J.1}$$

where $\Upsilon_l = 0.5$ if the direction of the traced straight line corresponds to the preferred direction of the lth actor cell, and is 0 otherwise. When the agent collides with the wall it takes a step in the opposite direction of length $|\delta(t)|$.

Other parameters are as in Appendix I except for: $R_{goal} = 500$ [reward units], $\tau_r = 3000$ ms, $\kappa = \tau_r/3$ for the TD error, the strength of lateral connectivity $\epsilon_1 = 1/28$ mV $\times$ ms; the learning rates $\eta_A = 10$ [reward units]$^{-1}$mV^{-1} and

$\eta_C = 15$ [reward units]$^{-1}$mV^{-1}. There is an increase in the speed, relative to the square maze problem, $s = 20$ cm/s and a decrease in the noise $\sigma_{n_0} = 0.5$.

K

EXPERIMENTAL PROTOCOL USED FOR FIGURE 5.22

Table K.1 shows the selected escape platforms for each day and the starting positions for each trial. The escape platforms and starting positions are shown in Figure K.1.

Day	Platform	Starting Position			
		Trial 1	Trial 2	Trial 3	Trial 4
1	2	E	W	S	N
2	5	W	N	E	S
3	4	N	S	W	E
4	7	S	E	W	N
5	3	E	S	W	N
6	6	S	N	W	E
7	8	W	E	N	S
8	1	N	W	S	E

Table K.1. Table indicating the starting position and goal location for each trial across the days.

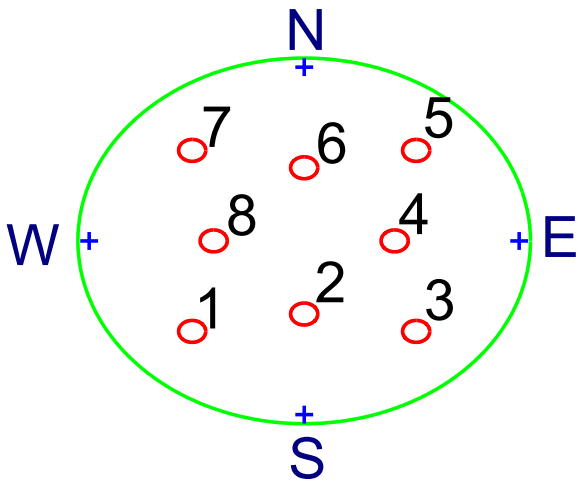


Figure K.1. Watermaze showing the platform (red circles) numbers and the names of the starting positions (blue crosses).

BIBLIOGRAPHY

[1] Merriam-Webster.com, "Neuroscience." https://www.merriam-webster.com/dictionary/neuroscience, May 2017.

[2] F. Gabbiani and S. J. Cox, *Mathematics for Neuroscientists*. Academic Press, 2010.

[3] Computational Neuroscience , University of Washington, "What is computational neuroscience?." http://compneuro.washington.edu/about/what-is-compneuro/, May 2017.

[4] P. Dayan and L. F. Abbott, *Theoretical Neuroscience*, vol. 806. Cambridge, MA: MIT Press, 2001.

[5] S. Seung, "Introduction to computational neuroscience [lecture notes]." Retrieved from https://ocw.mit.edu/courses/brain-and-cognitive-sciences/9-29j-introduction-to-computational-neuroscience-spring-2004/lecture-notes/lecture01.pdf, February 2013.

[6] A. L. Hodgkin and A. F. Huxley, "A quantitative description of membrane current and its application to conduction and excitation in nerve," *The Journal of physiology*, vol. 117, no. 4, p. 500, 1952.

[7] G. B. Ermentrout and D. H. Terman, *Mathematical foundations of neuroscience*, vol. 35. Springer Science & Business Media, 2010.

[8] S. Coombes, P. beim Graben, R. Potthast, and J. Wright, *Neural Fields*. Springer, 2014.

[9] E. M. Izhikevich, *Dynamical systems in neuroscience*. MIT press, 2007.

[10] D. Marr, *Vision: A computational investigation into the human representation and processing of visual information*. W.H. Freeman and Company, 1982.

[11] A. D. Redish, *Beyond the cognitive map: from place cells to episodic memory*. MIT Press, 1999.

[12] NC3RS, "National centre for the replacement, refinement & reduction of animals in research." http://www.nc3rs.org.uk/, May 2017.

[13] R. Llinas, "Neuron," *Scholarpedia*, vol. 3, no. 8, p. 1490, 2008. revision #91570.

[14] E. Niebur, "Electrical properties of cell membranes," *Scholarpedia*, vol. 3, no. 6, p. 7166, 2008. revision # 139387.

[15] S. Coombes, Y. Timofeeva, C.-M. Svensson, G. J. Lord, K. Josić, S. J. Cox, and C. M. Colbert, "Branching dendrites with resonant membrane: a "sum-over-trips" approach," *Biological Cybernetics*, vol. 97, no. 2, pp. 137–149, 2007.

[16] J. O'Keefe and M. L. Recce, "Phase relationship between hippocampal place units and the eeg theta rhythm," *Hippocampus*, vol. 3, no. 3, pp. 317–330, 1993.

[17] L. M. Giocomo, M.-B. Moser, and E. I. Moser, "Computational models of grid cells," *Neuron*, vol. 71, no. 4, pp. 589–603, 2011.

[18] T. Trappenberg, *Fundamentals of computational neuroscience*. OUP Oxford, 2002.

[19] J. J. Knierim and K. Zhang, "Attractor dynamics of spatially correlated neural activity in the limbic system," *Annual review of neuroscience*, vol. 35, pp. 267–285, 2012.

[20] R. U. Muller, J. B. Ranck, and J. S. Taube, "Head direction cells: properties and functional significance," *Current opinion in neurobiology*, vol. 6, no. 2, pp. 196–206, 1996.

[21] B. McNaughton, L. Chen, and E. Markus, ""dead reckoning," landmark learning, and the sense of direction: a neurophysiological and computational hypothesis," *Journal of Cognitive Neuroscience*, vol. 3, no. 2, pp. 190–202, 1991.

[22] T. Hafting, M. Fyhn, S. Molden, M.-B. Moser, and E. I. Moser, "Microstructure of a spatial map in the entorhinal cortex," *Nature*, vol. 436, no. 7052, pp. 801–806, 2005.

[23] C. N. Boccara, F. Sargolini, V. H. Thoresen, T. Solstad, M. P. Witter, E. I. Moser, and M.-B. Moser, "Grid cells in pre-and parasubiculum," *Nature neuroscience*, vol. 13, no. 8, pp. 987–994, 2010.

[24] E. I. Moser, Y. Roudi, M. P. Witter, C. Kentros, T. Bonhoeffer, and M.-B. Moser, "Grid cells and cortical representation," *Nature reviews. Neuroscience*, vol. 15, no. 7, p. 466, 2014.

[25] M. Witter, "Entorhinal cortex," *Scholarpedia*, vol. 6, no. 10, p. 4380, 2011. revision #90504.

[26] J. J. Couey, A. Witoelar, S.-J. Zhang, K. Zheng, J. Ye, B. Dunn, R. Czajkowski, M.-B. Moser, E. I. Moser, Y. Roudi, and M. P. Witter, "Recurrent inhibitory circuitry as a mechanism for grid formation," *Nature neuroscience*, vol. 16, no. 3, pp. 318–324, 2013.

[27] D. Bush, C. Barry, and N. Burgess, "What do grid cells contribute to place cell firing?," *Trends in Neurosciences*, 2014.

[28] N. Burgess, C. Barry, K. J. Jeffery, and J. O'Keefe, "A grid & place cell model of path integration utilizing phase precession versus theta," *Trial*, vol. 1, no. 1.7, pp. 3–5, 2005.

[29] M. C. Fuhs and D. S. Touretzky, "A spin glass model of path integration in rat medial entorhinal cortex," *The Journal of Neuroscience*, vol. 26, no. 16, pp. 4266–4276, 2006.

[30] B. L. McNaughton, F. P. Battaglia, O. Jensen, E. I. Moser, and M.-B. Moser, "Path integration and the neural basis of the 'cognitive map'," *Nature Reviews Neuroscience*, vol. 7, no. 8, pp. 663–678, 2006.

[31] H. Mhatre, A. Gorchetchnikov, and S. Grossberg, "Grid cell hexagonal patterns formed by fast self-organized learning within entorhinal cortex," *Hippocampus*, vol. 22, no. 2, pp. 320–334, 2012.

[32] P. K. Pilly and S. Grossberg, "Spiking neurons in a hierarchical self-organizing map model can learn to develop spatial and temporal properties of entorhinal grid cells and hippocampal place cells," *PloS one*, vol. 8, no. 4, p. e60599, 2013.

[33] S. Grossberg and P. K. Pilly, "How does the modular organization of entorhinal grid cells develop?," *Frontiers in Human Neuroscience*, vol. 8, p. 337, 2014.

[34] E. Kropff and A. Treves, "The emergence of grid cells: Intelligent design or just adaptation?," *Hippocampus*, vol. 18, no. 12, pp. 1256–1269, 2008.

[35] Y. Dordek, D. Soudry, R. Meir, and D. Derdikman, "Extracting grid cell characteristics from place cell inputs using non-negative principal component analysis," *Elife*, vol. 5, p. e10094, 2016.

[36] T. Bonnevie, B. Dunn, M. Fyhn, T. Hafting, D. Derdikman, J. L. Kubie, Y. Roudi, E. I. Moser, and M.-B. Moser, "Grid cells require excitatory drive from the hippocampus," *Nature neuroscience*, vol. 16, no. 3, pp. 309–317, 2013.

[37] R. F. Langston, J. A. Ainge, J. J. Couey, C. B. Canto, T. L. Bjerknes, M. P. Witter, E. I. Moser, and M.-B. Moser, "Development of the spatial

representation system in the rat," *Science*, vol. 328, no. 5985, pp. 1576–1580, 2010.

[38] K. B. Kjelstrup, T. Solstad, V. H. Brun, T. Hafting, S. Leutgeb, M. P. Witter, E. I. Moser, and M.-B. Moser, "Finite scale of spatial representation in the hippocampus," *Science*, vol. 321, no. 5885, pp. 140–143, 2008.

[39] Y. Burak and I. R. Fiete, "Accurate path integration in continuous attractor network models of grid cells," *PLoS computational biology*, vol. 5, no. 2, p. e1000291, 2009.

[40] K. Yoon, M. A. Buice, C. Barry, R. Hayman, N. Burgess, and I. R. Fiete, "Specific evidence of low-dimensional continuous attractor dynamics in grid cells," *Nature neuroscience*, vol. 16, no. 8, pp. 1077–1084, 2013.

[41] P. Beed, A. Gundlfinger, S. Schneiderbauer, J. Song, C. Böhm, A. Burgalossi, M. Brecht, I. Vida, and D. Schmitz, "Inhibitory gradient along the dorsoventral axis in the medial entorhinal cortex," *Neuron*, vol. 79, no. 6, pp. 1197–1207, 2013.

[42] B. Ermentrout, "Stripes or spots? nonlinear effects in bifurcation of reaction-diffusion equations on the square," in *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, vol. 434, pp. 413–417, The Royal Society, 1991.

[43] K. Zhang, "Representation of spatial orientation by the intrinsic dynamics of the head-direction cell ensemble: a theory," *The journal of neuroscience*, vol. 16, no. 6, pp. 2112–2126, 1996.

[44] R. Curtu and B. Ermentrout, "Pattern formation in a network of excitatory and inhibitory cells with adaptation," *SIAM Journal on Applied Dynamical Systems*, vol. 3, no. 3, pp. 191–231, 2004.

[45] A. Roxin, N. Brunel, and D. Hansel, "Role of delays in shaping spatiotemporal dynamics of neuronal activity in large networks," *Physical review letters*, vol. 94, no. 23, p. 238103, 2005.

[46] L. M. Giocomo, E. A. Zilli, E. Fransén, and M. E. Hasselmo, "Temporal frequency of subthreshold oscillations scales with entorhinal grid cell field spacing," *Science*, vol. 315, no. 5819, pp. 1719–1722, 2007.

[47] P. Bressloff and S. Coombes, "Physics of the extended neuron," *International Journal of Modern Physics B*, vol. 11, no. 20, pp. 2343–2392, 1997.

[48] Z. F. Mainen and T. J. Sejnowski, "Influence of dendritic structure on firing pattern in model neocortical neurons," *Nature*, vol. 382, no. 6589, p. 363, 1996.

[49] M. Leutenegger, "Error function of complex numbers." https://uk.mathworks.com/matlabcentral/fileexchange/18312-error-function-of-complex-numbers, 2008.

[50] L. M. Giocomo and M. E. Hasselmo, "Time constants of h current in layer II stellate cells differ along the dorsal to ventral axis of medial entorhinal cortex," *Journal of Neuroscience*, vol. 28, no. 38, pp. 9414–9425, 2008.

[51] M. E. Hasselmo, "Neuronal rebound spiking, resonance frequency and theta cycle skipping may contribute to grid cell firing in medial entorhinal cortex," *Philosophical Transactions Royal Society Biological Sciences*, vol. 369, no. 1635, p. 20120523, 2014.

[52] R. Dodla, *Encyclopedia of Computational Neuroscience*, ch. Postinhibitory Rebound and Facilitation, pp. 1–3. New York, NY: Springer New York, 2013.

[53] N. Burgess, C. Barry, and J. O'Keefe, "An oscillatory interference model of grid cell firing," *Hippocampus*, vol. 17, no. 9, pp. 801–812, 2007.

[54] E. T. Reifenstein, R. Kempter, S. Schreiber, M. B. Stemmler, and A. V. Herz, "Grid cells in rat entorhinal cortex encode physical space with independent firing fields and phase precession at the single-trial level," *Proceedings of the National Academy of Sciences*, vol. 109, no. 16, pp. 6301–6306, 2012.

[55] M. F. Nolan, J. T. Dudman, P. D. Dodson, and B. Santoro, "HCN1 channels control resting and active integrative properties of stellate cells from layer II of the entorhinal cortex," *Journal of Neuroscience*, vol. 27, no. 46, pp. 12440–12451, 2007.

[56] Z. Navratilova, L. M. Giocomo, J.-M. Fellous, M. E. Hasselmo, and B. L. McNaughton, "Phase precession and variable spatial scaling in a periodic attractor map model of medial entorhinal grid cells with realistic after-spike dynamics," *Hippocampus*, vol. 22, no. 4, pp. 772–789, 2012.

[57] M. E. Hasselmo and C. F. Shay, "Grid cell firing patterns may arise from feedback interaction between intrinsic rebound spiking and transverse traveling waves with multiple heading angles," *Frontiers in systems neuroscience*, vol. 8, p. 201, 2014.

[58] Y. Tsuno, G. W. Chapman, and M. E. Hasselmo, "Rebound spiking properties of mouse medial entorhinal cortex neurons in vivo," *European Journal of Neuroscience*, vol. 42, no. 11, pp. 2974–2984, 2015.

[59] C. F. Shay, M. Ferrante, G. W. Chapman, and M. E. Hasselmo, "Rebound spiking in layer II medial entorhinal cortex stellate cells: Possible mechanism of grid cell function," *Neurobiology of learning and memory*, vol. 129, pp. 83–98, 2016.

[60] M. Ferrante, C. F. Shay, Y. Tsuno, G. W. Chapman, and M. E. Hasselmo, "Post-inhibitory rebound spikes in rat medial entorhinal layer II/III principal cells: in vivo, in vitro, and computational modeling characterization," *Cerebral Cortex*, vol. 27, no. 3, pp. 2111–2125, 2017.

[61] N. Kopell and G. LeMasson, "Rhythmogenesis, amplitude modulation, and multiplexing in a cortical architecture," *Proceedings of the National Academy of Sciences*, vol. 91, no. 22, pp. 10586–10590, 1994.

[62] J. G. Heys and M. E. Hasselmo, "Neuromodulation of I_h in layer II medial entorhinal cortex stellate cells: a voltage-clamp study," *Journal of Neuroscience*, vol. 32, no. 26, pp. 9066–9072, 2012.

[63] C. F. Shay, I. S. Boardman, N. M. James, and M. E. Hasselmo, "Voltage dependence of subthreshold resonance frequency in layer II of medial entorhinal cortex," *Hippocampus*, vol. 22, no. 8, pp. 1733–1749, 2012.

[64] R. N. Bracewell, *The Fourier transform and its applications*, vol. 31999. McGraw-Hill New York, 1986.

[65] J. W. Evans, "The stable and the unstable impulse," *Indiana University Mathematics Journal*, vol. 24, no. 12, 1975.

[66] S. Coombes and M. R. Owen, "Evans functions for integral neural field equations with Heaviside firing rate function," *SIAM Journal on Applied Dynamical Systems*, vol. 3, no. 4, pp. 574–600, 2004.

[67] S. Coombes, R. Thul, and K. C. Wedgwood, "Nonsmooth dynamics in spiking neuron models," *Physica D: Nonlinear Phenomena*, vol. 241, no. 22, pp. 2042–2057, 2012.

[68] F. Sargolini, M. Fyhn, T. Hafting, B. L. McNaughton, M. P. Witter, M.-B. Moser, and E. I. Moser, "Conjunctive representation of position, direction, and velocity in entorhinal cortex," *Science*, vol. 312, no. 5774, pp. 758–762, 2006.

[69] M. Bonilla-Quintana, K. C. Wedgwood, R. D. O'Dea, and S. Coombes, "An analysis of waves underlying grid cell firing in the medial enthorinal cortex," *The Journal of Mathematical Neuroscience*, vol. 7, no. 1, p. 9, 2017.

[70] T. Hartley, C. Lever, N. Burgess, and J. O'Keefe, "Space in the brain: how the hippocampal formation supports spatial cognition," *Phil. Trans. R. Soc. B*, vol. 369, no. 1635, p. 20120510, 2014.

[71] I. R. Fiete, Y. Burak, and T. Brookings, "What grid cells convey about rat location," *Journal of Neuroscience*, vol. 28, no. 27, pp. 6858–6871, 2008.

[72] H. Stensola, T. Stensola, T. Solstad, K. Frøland, M.-B. Moser, and E. I. Moser, "The entorhinal grid map is discretized," *Nature*, vol. 492, no. 7427, pp. 72–78, 2012.

[73] D. C. Rowland, Y. Roudi, M.-B. Moser, and E. I. Moser, "Ten years of grid cells," *Annual review of neuroscience*, vol. 39, pp. 19–40, 2016.

[74] X.-X. Wei, J. Prentice, and V. Balasubramanian, "A principle of economy predicts the functional architecture of grid cells," *ELife*, vol. 4, p. e08362, 2015.

[75] M. Witter, H. Groenewegen, F. L. Da Silva, and A. Lohman, "Functional organization of the extrinsic and intrinsic circuitry of the parahippocampal region," *Progress in neurobiology*, vol. 33, no. 3, pp. 161–253, 1989.

[76] S. Cheng and L. M. Frank, "The structure of networks that produce the transformation from grid cells to place cells," *Neuroscience*, vol. 197, pp. 293–306, 2011.

[77] U. M. Erdem and M. Hasselmo, "A goal-directed spatial navigation model using forward trajectory planning based on grid cells," *European Journal of Neuroscience*, vol. 35, no. 6, pp. 916–931, 2012.

[78] B. Poucet, F. Sargolini, E. Y. Song, B. Hangya, S. Fox, and R. U. Muller, "Independence of landmark and self-motion-guided navigation: a different role for grid cells," *Philosophical Transactions Royal Society Biological Sciences*, vol. 369, no. 1635, p. 20130370, 2014.

[79] D. Foster, R. Morris, P. Dayan, *et al.*, "A model of hippocampally dependent navigation, using the temporal difference learning rule," *Hippocampus*, vol. 10, no. 1, pp. 1–16, 2000.

[80] E. C. Tolman, B. Ritchie, and D. Kalish, "Studies in spatial learning. ii. place learning versus response learning.," *Journal of Experimental Psychology*, vol. 36, no. 3, p. 221, 1946.

[81] E. C. Tolman, B. F. Ritchie, and D. Kalish, "Studies in spatial learning. i. orientation and the short-cut.," *Journal of Experimental Psychology*, vol. 36, no. 1, p. 13, 1946.

[82] J. O'Keefe and J. Dostrovsky, "The hippocampus as a spatial map. preliminary evidence from unit activity in the freely-moving rat," *Brain research*, vol. 34, no. 1, pp. 171–175, 1971.

[83] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*, vol. 1. MIT press Cambridge, 1998.

[84] J. O. Keefe and N. Burgess, "Geometric determinants of the place fields of hippocampal neurons," *Nature*, vol. 381, no. 6581, p. 425, 1996.

[85] K. Doya, "Reinforcement learning in continuous time and space," *Neural computation*, vol. 12, no. 1, pp. 219–245, 2000.

[86] R. Bellman, *Dynamic Programming*. Princeton University Press, 1957.

[87] V. Gullapalli, "A stochastic reinforcement learning algorithm for learning real-valued functions," *Neural networks*, vol. 3, no. 6, pp. 671–692, 1990.

[88] N. Frémaux, H. Sprekeler, and W. Gerstner, "Reinforcement learning using a continuous time actor-critic framework with spiking neurons," *PLoS Comput Biol*, vol. 9, no. 4, p. e1003024, 2013.

[89] K. C. Berridge and T. E. Robinson, "Parsing reward," *Trends in neurosciences*, vol. 26, no. 9, pp. 507–513, 2003.

[90] T. Bast, I. A. Wilson, M. P. Witter, and R. G. Morris, "From rapid place learning to behavioral performance: a key role for the intermediate hippocampus," *PLoS Biol*, vol. 7, no. 4, p. e1000089, 2009.

[91] F. Savelli, D. Yoganarasimha, and J. J. Knierim, "Influence of boundary removal on the spatial representations of the medial entorhinal cortex," *Hippocampus*, vol. 18, no. 12, pp. 1270–1282, 2008.

[92] Q. J. Huys, R. Cools, M. Gölzer, E. Friedel, A. Heinz, R. J. Dolan, and P. Dayan, "Disentangling the roles of approach, activation and valence in instrumental and pavlovian responding," *PLoS computational biology*, vol. 7, no. 4, p. e1002028, 2011.

[93] T. Poggio and F. Girosi, "A theory of networks for approximation and learning," Tech. Rep. No. AI-M-1140, Massachusetts Institute of Technology, Cambridge, Artificial Intelligence Lab, 1989.

[94] D. Terman, "Stability of planar wave solutions to a combustion model," *SIAM journal on mathematical analysis*, vol. 21, no. 5, pp. 1139–1171, 1990.

[95] K. J. Jeffery, H. J. Page, and S. M. Stringer, "Optimal cue combination and landmark-stability learning in the head direction system," *The Journal of Physiology*, vol. 594, no. 22, pp. 6527–6534, 2016.

[96] W.-h. Zhang, A. Chen, M. J. Rasch, and S. Wu, "Decentralized multisensory information integration in neural systems," *Journal of Neuroscience*, vol. 36, no. 2, pp. 532–547, 2016.

[97] S. Coombes, "Dynamics of synaptically coupled integrate-and-fire-or-burst neurons," *Physical Review E*, vol. 67, no. 4, p. 041910, 2003.